Coherent Diffraction Imaging with Low-Energy Electrons

Dissertation
zur
Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)

vorgelegt der
Mathematisch-naturwissenschaftlichen Fakultät
der
Universität Zürich

von
Elvira Steinwand
aus
Deutschland

Promotionskomitee
Prof. Dr. Hans-Werner Fink (Vorsitz und Leitung der Dissertation)
Dr. Heinz Gross

Zürich, 2010

Gutachter:

Dr. Roger Morin
CNRS, Marseille

Abstract

In the work presented here, a novel method for structural analysis on individual molecules is introduced. Once established, the technique shall significantly advance the field of structural biology, since averaging over millions of molecules, a requisite for conventional methods, will be obsolete. In particular, the investigation of proteins that do not readily crystallize shall thus become possible. However, for atomic resolution, such an imaging method requires both, radiation with wavelengths in the Ångstrom regime and the possibility to acquire sufficient data from a single molecule without destroying it. It has been shown that low-energy electrons, i.e. electrons with kinetic energies between 50 and 250 eV, fulfil these requirements.

Here, an experimental scheme is presented, devised for coherent diffraction imaging of individual molecules with low-energy electrons. In coherent diffraction imaging, a plane wave is scattered by the sample, and the far-field diffraction pattern is recorded. Under certain conditions, the phases of the diffraction pattern can be retrieved by numerical algorithms and with this, an image of the sample can be reconstructed.

A pivotal component to realize such an experiment with low-energy electrons is an electrostatic microlens for collimating the divergent electron beam emitted by the source. Several methods developed for the fabrication of such microlenses are presented here, and it is shown, that the reduced lens size leads to reduced spherical aberrations. The use of a low-aberration microlens in combination with an electron source of atomic dimension emission area, makes it possible to minimize the divergence angle of the beam to a few milliradians. This allows coherent illumination of an object of a few hundred nanometers in size placed at about 200 µm beyond the lens.

With regard to a proper phase retrieval, it is essential to expose the sample as bare as possible to the electron beam because any additional substrate affects the diffraction pattern. Hence, the molecules are suspended across holes in a support film. First experiments have been carried out using carbon nanotubes as samples, also due to the fact that they might serve as templates for protein sample preparation in the future. The preparation methods for placing carbon nanotubes across hole- or slit-structures in carbon films are described.

First diffraction patterns of carbon nanotube structures could be recorded. They show intensity modulations up to the seventh order perpendicular to the longitudinal direction of the structures and the distance between the modulation maxima can be associated to the width of the structures. However, the faint speckles in the diffraction patterns, containing more detailed information about the shape of the objects, could not yet be resolved. Hence, subsequent numerical image reconstruction was not possible so far. This is mainly due to the limited dynamic and spatial resolution of the detector, whereas the spatial coherence of the incident electrons currently is not the limiting factor.

Zusammenfassung

In der vorliegenden Arbeit soll eine neuartige Methode zur Strukturanalyse vorgestellt werden, deren Ziel es ist, Strukturinformation anhand der Untersuchung eines einzelnen Moleküls zu erhalten. Für das Gebiet der Strukturbiologie würde diese Methode einen bedeutenden Fortschritt darstellen, da das Mitteln über Millionen von Molekülen – bei konventionellen Methoden eine Notwendigkeit – hinfällig würde. Insbesondere könnten so auch Proteine untersucht werden, die sich nur schwer oder überhaupt nicht kristallisieren lassen. Im Hinblick auf atomare Auflösung erfordert solch eine Abbildungsmethode jedoch Strahlung mit einer Wellenlänge im Ångstrombereich, die es zudem erlaubt eine genügend grosse Anzahl elastischer Streuereignisse zu detektieren, ohne dass das Molekül währenddessen zerstört wird. Es wurde gezeigt, dass niederenergetische Elektronen, d.h. Elektronen mit einer kinetischen Energie zwischen 50 und 250 eV, diese Anforderungen erfüllen.

In der vorliegenden Arbeit wird der experimentelle Aufbau beschrieben, der die Abbildung einzelner Moleküle mittels Beugung kohärenter niederenergetischer Elektronen ermöglichen soll. Bei der bildgebenden Methode der kohärenten Beugung wird eine ebene Welle an einem Objekt gestreut und das Beugungsbild im Fernfeld detektiert. Sind bestimmte Voraussetzungen erfüllt, so können die Phasen des Beugungsbildes mittels numerischer Algorithmen wiedergewonnen werden, womit man schliesslich ein Bild des Objekts erhält.

Eine Schlüsselkomponente für die Realisierung eines solchen Experiments mit niederenergetischen Elektronen ist eine elektrostatische Mikrolinse, die dazu dient, den von der Quelle emittierten divergenten Elektronenstrahl zu kollimieren. Die für die Herstellung einer solchen Mikrolinse entwickelten Methoden werden beschrieben, und es wird gezeigt, dass die reduzierte Grösse der Linsen zu verringerten sphärischen Aberrationen führt. Eine Linse mit geringen sphärischen Aberrationen in Kombination mit einer Elektronenquelle, die eine effektive Emissionsregion atomarer Grösse hat, ermöglicht es, den Divergenzwinkel des Strahls auf wenige Milliradian zu minimieren. Dadurch kann eine Probe von einigen hundert Nanometern Grösse, die etwa 200 µm hinter der Linse platziert wird, kohärent beleuchtet werden.

Im Hinblick auf die anschliessende Phasenrekonstruktion sollte die Probe möglichst freitragend dem Elektronenstrahl ausgesetzt sein, da jegliche Streuung an einem Substrat das Beugungsbild beeinträchtigt. Daher werden die Moleküle über Löcher in einem Probenträgerfilm gelegt. Für erste Experimente wurden Kohlenstoff-Nanoröhrchen als Testobjekte verwendet, nicht zuletzt weil sie für die künftige Präparation von Proteinen als eine Art Probenträger dienen könnten. Methoden, Kohlenstoff-Nanoröhrchen so zu präparieren, dass sie über Loch- oder Schlitzstrukturen in Kohlenstofffilmen zu liegen kommen, werden hier vorgestellt.

Die ersten Beugungsbilder von Kohlenstoff-Nanoröhrchen, die mit dem hier präsentierten experimentellen Aufbau aufgenommen werden konnten, zeigen Intensitätsmodulationen senkrecht zur Längsrichtung der Nanoröhrchen bis zur siebten Beugungsordnung. Dabei entspricht der Abstand zwischen den Modulationsmaxima dem Durchmesser der Strukturen. Die schwachen Speckles des Beugungsbildes, welche detaillierte Information über die Form des Objekts enthielten, konnten bis anhin allerdings noch nicht aufgelöst werden, weswegen eine anschliessende numerische

Phasenrekonstruktion nicht möglich war. Das liegt vor allem an der begrenzten Dynamik und Auflösung des Detektors, während die räumliche Kohärenz der einfallenden Elektronen derzeit nicht limitierend ist.

Contents

List of Acronyms

AFM	atomic force microscopy
BHF	buffered hydrofluoric acid
CCD	charged coupled device
CDI	coherent diffraction imaging
CNT	carbon nanotube
cryo-EM	cryo-electron microscopy
DFT	discrete Fourier transform
ER-algorithm	error-reduction algorithm
FIB	focused ion beam
FOP	fibre optic plate
FT	Fourier transform
FWHM	full width at half maximum
HIO	hybrid input-output
LEEPS	low-energy electron point source microscopy
MCP	microchannel plate
NMR	nuclear magnetic resonance spectroscopy
SEM	scanning electron microscopy
SiN	silicon nitride
TEM	transmission electron microscopy
UHV	ultra high vacuum
XFEL	x-ray free electron laser

Chapter 1

Introduction

It is very easy to answer many
of the fundamental biological questions;
you just look at the thing!
Richard P. Feynman

The wish to understand the molecular mechanisms underlying biological processes in a living cell has driven the research in structural biology during the past decades. Revealing the structure of a biological molecule may help to figure out its function in the complicated interaction network within a cell. Several methods for structure determination of proteins are nowadays established, having led to an impressive database of protein structures that are known with atomic resolution. Most of these structures have been determined by x-ray crystallography, since this method is superior in terms of achievable resolution. However, other methods namely nuclear magnetic resonance (NMR) spectroscopy and cryo electron microscopy (cryo-EM) have also contributed significantly to the substantial progress in this field. The different methods can often complement each other concerning their resolution and applicability. Imaging protein structures is far more intricate than *just looking at it and see where the atoms are* [1]. It is rather an iterative process of data collection, structure solution and model refinement, putting together the puzzle pieces of information that have been obtained by different methods.

The mentioned established methods for protein structure determination have all in common, that they involve averaging over a large number of molecules instead of looking at one single molecule. Since most of the proteins may exhibit different conformations, fine details of their structure may be smeared out by averaging. In NMR, where the proteins are dissolved in water, this is the main resolution limiting factor, especially concerning the more flexible parts toward the molecular surface [2]. Having the proteins in an aqueous solution, on the other hand, is one of the advantages of NMR, since this at least resembles their natural environment despite the relatively high concentrations necessary. The structure of a protein forced into a crystal in contrast, may deviate significantly from its natural conformation. While x-ray crystallography can provide the structure of proteins with atomic resolution, the purification of proteins and subsequent crystallization is an often fairly time consuming tedious task. Moreover, there are about 10^6 human proteins that are impossible or very difficult to crystallize, including also a large number of membrane proteins [3]. Since the latter are the ones responsible for the interaction of the cell with its environment, they are of outstanding importance for drug design. NMR in turn is only applicable to small proteins up to a size of about 15 kDa [2]. Hence, there is a strong demand for alternative methods for determining the atomic structure of proteins without the need to crystallize them. Cryo-EM has evolved during the past two decades as a technique for imaging proteins and other biological macromolecules, lifting some of the restrictions of x-ray crystallography and NMR. Having advantages in view of sample preparation, including the fact that the molecular structure is preserved by vitrification and cooling, obtaining structural information from just one single molecule is not possible by cryo-EM either. The maximal dose of high-energy electrons that can be applied to a biomolecule at liquid helium temperatures before being destroyed by the radiation amounts to only $5 - 10\,\text{electrons}/\,\text{Å}^2$, resulting in extremely noisy images. The signal to noise ratio thus has to be improved by averaging over a large number of molecules, usually in the order of 10^4 [4]. Therefore, cryo-EM has initially been applied to two-dimensional crystalline structures of proteins [5], but it is now routinely applied also to uncrystallized macromolecules randomly orientated in a vitrified solution. As in crystallography, a lot of data processing is required to extract structural information about the molecules from the large number of recorded images, including algorithms for alignment and angle assignment. Although resolutions well below 10 Å could be achieved in some cases [6], the resolution of cryo-EM data is usually in the range of 10 to 20 Å [4] and, hence, such data alone do in general not suffice for the building of atomic models. They are often used complementing data obtained with x-ray crystallography. Resolution in cryo-EM is not limited by instrumental resolution of the microscope, but by the maximal exposure dose and by the required averaging over thousands of molecules.

The development of the x-ray free-electron lasers (XFEL) raised hope to obtain structural information from a single biomolecule and thus to circumvent the problems going along with averaging over large numbers of molecules. The extremely short pulse length of the XFELs of 10 fs and their high brightness shall allow capturing an image of the molecule before it disassembles due to inelastic scattering processes [7]. However, as has been shown in further calculations, the relation between inelastic and elastic scattering processes is so unfavourable for x-rays with a wavelength of 1 Å, that even the high brightness of an XFEL source of 10^{12} photons per pulse, would not suffice to gather enough information from one shot [8]. This means that averaging over many copies of the molecule is also required in this case including the same issues of orientation assignment and possible conformational changes as in cryo-EM.

It appears that low-energy electrons with kinetic energies between 60 and 250 eV are the only known radiation today with wavelengths in the Ångstrom range where elastic scattering dominates destructive inelastic processes [9]. This in turn allows to obtain the information about the atomic structure of a molecule from one single copy, with no need for averaging. One can think of different ways how to collect this information. A technique already existing for several years is low-energy electron holography [10], according to the scheme of in-line holography initially proposed by Gabor [11]. A spherical electron wave is scattered by the object. At a distant screen, the unscattered part of the wave and the scattered part are superposed. As the unscattered reference wave is known, the phases of the scattered wave are encoded in the recorded amplitude of the hologram and can be retrieved by numerical means.

An alternative approach for the imaging of individual biological molecules with low-energy electrons, which is the subject of the present work, is to direct a parallel beam of low-energy electrons towards the molecule and record the far-field diffraction pattern at a distant detector. An image of the object can then be constructed numerically, with the computer performing the task that is done by a lens in a conventional microscope. This is the principle already applied in x-ray experiments for imaging of non-crystalline specimens and conceived for imaging with the XFELs. Numerical reconstruction requires high spatial coherence of the initial beam. This is why the method is also referred to as *coherent* diffraction imaging (CDI).

The experimental realization of a coherent diffraction experiment with low-energy electrons requires several dedicated components, the most challenging of them being an electron lens collimating the divergent electron beam, as it is emitted by the source, without destroying spatial coherence. For this purpose, an electrostatic lens of micrometer dimensions has been developed, miniaturized in order to reduce aberrations. This microlens is the subject of Part I of the work presented here. After a brief introduction the fabrication methods for microlenses will be described in Chapter 3. How the produced lenses have been tested and characterized in a quantitative way is explained in Chapter 4. Part I concludes with a summary of the obtained results and an outlook on possible applications for the microlens beside its implementation in a low-energy electron diffraction experiment.

Part II of this thesis is about coherent low-energy electron diffraction at individual molecules. After introducing the principle of coherent diffraction imaging in Chapter 6, some advantages of low-energy electrons in respect of their interaction with matter and radiation damage will be pointed out and a concept for the description of the scattering process will be given in Chapter 7. Chapter 8 will analyse the issue of the phase problem, always associated with the recording of diffraction patterns, including a survey of the existing algorithms for phase retrieval in CDI. The intrinsic limitations on the achievable resolution by spatial and temporal coherence will be discussed in Chapter 9. The overall setup of the low-energy electron coherent diffraction experiment will be presented in Chapter 10, describing and characterizing in detail the particular components and the methods for sample preparation. First diffraction patterns that could be obtained during the course of the work presented are shown in Chapter 11.

Part I

Microlens

Chapter 2

Down-scaling approach for reducing spherical aberrations

The spherical aberration has the effect of focusing electron rays passing through outer zones of a lens at a shorter distance than paraxial rays. This leads to an extended focal spot as schematically illustrated in Fig. 2.1. Spherical aberrations are much more severe in electron optics than for visible light where they can be corrected by the use of aspherical lenses or by combination of several lenses with opposite amounts of spherical aberration. As has been recognized by Scherzer in 1936 [12], the coefficients for spherical as well as for chromatic aberrations are always positive for a stationary, rotationally symmetric electron lens without space-charges; hence, spherical aberrations are unavoidable with these lenses. Their effect can be mitigated by the use of apertures, cutting off the electron rays far from the optical axis and allowing only a small acceptance angle. Correction of spherical aberration in electron lenses however is only possible by lifting any one of the constraints of the Scherzer theorem, hence by introducing time-varying fields or space charges or by abandoning rotational symmetry [13]. This latter approach has been pursued by Krivanek et al. who first succeeded in actually improving the resolution of a scanning transmission electron microscope (STEM) by implementation of a quadrupole-octupole corrector for reduction of spherical aberrations [14, 15]. This breakthrough revolutionized electron microscopy, finally enabling imaging at atomic resolution [16], and multipole correctors are now also implemented in commercial state-of-the-art electron microscopes.

A different approach towards minimal aberration in electron lenses relies on the concept of scaling down lens dimensions. In principle, this can be done in two different ways: Scaling all lens

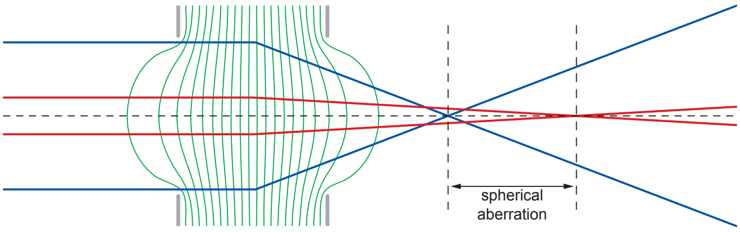


Figure 2.1: *Schematic illustration of the spherical aberrations of a two-electrode lens.*

dimensions uniformly by a constant factor k, while keeping the electrode potentials and the initial electron energy unchanged, results in electric field strengths scaled up by a factor 1/k. Hence, the velocity vector of a charged particle passing the lens is changed by a certain amount within a shorter distance due to increased field strength. The overall shape of the electron trajectories is thus maintained; they just undergo the same similarity transformation as the lens geometry. Alternatively, potentials and electron energies can also be scaled by the same factor k as the lens geometry, which obviously leads as well to a mere scaling of the electron trajectories while their shape is not affected. Since spherical aberration manifests itself in the size of the aberration disk, it thus scales directly with the lens dimensions in both cases.

In connection with an electron source as used for field emission of low-energy electrons, which exhibits an effective emission area of near atomic dimension, and thus can be regarded as a down-scaled version of a conventional field emitter, it suggests itself to use an electron-optical system with aberrations reduced by miniaturization as well. One advantage of combining a low-aberration lens with a miniaturized electron source is that there is no need to implement apertures for reducing the effect of spherical aberrations. A larger part of the electron current initially emitted by the source can hence be gathered by the lens, resulting in an enhanced beam brightness. Second, the combination of a small source size with low-aberration lens is advantageous if it comes to creating a focus. Since the focus is nothing but the image of the primary source, a small focal spot can be generated even without demagnification. The other way round, a source size of near atomic dimensions is useless, if generating a small focal spot is obstructed by lens aberrations. An electron-optical system consisting of a miniaturized electron source and miniaturized lenses, can thus be considered as a down-scaled version of a conventional electron column with a hairpin field emitter and macroscopic lenses, where potentials and energies are also scaled.

In principle, a miniaturized electron-optical system can either be based on electrostatic or on magnetic lenses. The latter are usually favoured for conventional systems, mainly because they exhibit lower spherical aberrations especially at short working distances compared to their electrostatic counterparts. However, electrostatic lenses appear to be the better candidates for miniaturization, since their compactness and simplicity makes them more compatible with micro-fabrication processes. While several attempts have been undertaken in the direction of miniaturization of electrostatic electron lenses in the past [17, 18, 19, 20, 21], routine operation of low-aberration micrometer scaled lenses and their application in scientific instruments is still lacking.

In connection with an experiment for coherent low-energy diffraction as it will be described in Part II of the work presented here, such a miniaturized electron-optical system shall enable the generation of an almost parallel beam of high brightness. Therefore, developing fabrication methods for electrostatic microlenses has been essential for the setup of the diffraction experiment.

In the following, several different fabrication methods will first be presented. As will be described subsequently, the lenses have been experimentally tested and their performance has been characterized also in a quantitative way by comparison of measured values of the diameter of the collimated beam with values obtained from ray tracing simulations. The following sections are partly excerpts of an article [22] which is also appended to this work in Appendix A.

Chapter 3

Microlens Fabrication

3.1 Principle of operation and basic design criteria

The simplest type of electrostatic lens consists of two parallel planar electrodes with two concentric apertures of not necessarily the same size. The electrodes must be separated by an insulating material to maintain a voltage between them, resulting in a focusing electrostatic field distribution at the apertures as illustrated in Fig. 3.1. A focusing effect is generated independent of the polarity of the applied lens-voltage, the electron trajectories however look different for the accelerating than for the decelerating mode. A lens size as small as possible appears preferable in order to minimize spherical aberrations according to the down-scaling approach. Considering established micro-fabrication techniques and available materials, lens dimensions in the range of one micrometer appear sensible. Micrometer lens dimensions shall also ensure easy positioning of the lens in an electron-optical system using conventional nano-positioning devices based on piezoelectric manipulators.

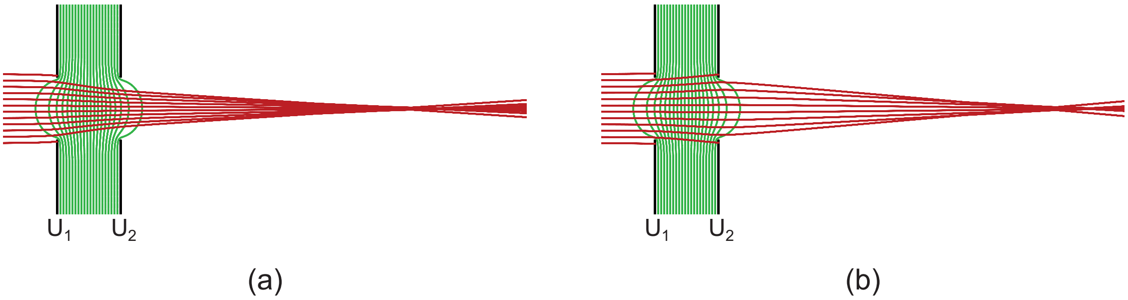


Figure 3.1: *Operating principle of an electrostatic two-electrode aperture lens (a) Accelerating electrical field between the electrodes. Equipotential lines and electron trajectories have been calculated for* $U_1 = 100\,\mathrm{V}$ *and* $U_2 = 350\,\mathrm{V}$ *for an initial electron energy of 100 eV. (b) Decelerating electrical field between the electrodes. Equipotential lines and electron trajectories have been calculated for* $U_1 = 100\,\mathrm{V}$ *and* $U_2 = 30\,\mathrm{V}$ *for an initial electron energy of 100 eV.*

3.2 Lens fabrication methods

We have developed several lens fabrication methods comprising various micro-fabrication steps, evaporation methods and materials for the insulating layers and lens electrodes. The result of all methods is a lens structure similar to the one depicted in Fig. 3.2 (a) featuring a 1 µm thick insulating layer, two electrodes of several ten nanometers thickness with micrometer sized apertures in them. The electrodes must be free-standing around the apertures to avoid charging effects of the insulating walls while the electron beam impinges on the lens. The main principle difference between the developed lens fabrication methods consists in the way of obtaining this free-standing region. For all types of microlenses commercially available silicon nitride (SiN) membranes[1] of one micrometer thickness as they are shown in Fig. 3.2 have been used as a starting material. They serve as insulating layers between the two lens electrodes and exhibit a measured breakdown voltage around 320 V, high enough to focus a beam of electrons with kinetic energies in the range of 100 eV no matter whether focusing is achieved by an accelerating or a decelerating voltage. The several fabrication steps for three different types of microlenses will be described in the following in more detail.

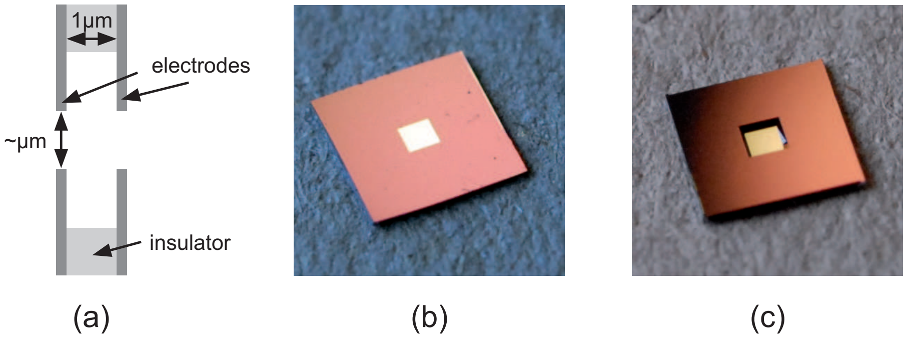


Figure 3.2: *(a) Basic design of a two-electrode lens with micrometer sized apertures. The electrodes are separated by an insulating layer of 1 µm in thickness. To avoid charging of the insulating walls by impinging electrons, the electrodes must be free-standing around the aperture. (b) and (c) Upper and bottom side of a SiN membrane. The membranes are available as 5 × 5 mm chips consisting of a silicon frame of 200 µm in thickness with a 1 × 1 mm window where the SiN membrane is free standing. On the bottom side of the chip the membrane is recessed relative to the frame.*

3.2.1 Lens fabrication by back-etching

One possibility to achieve that the electrodes are free-standing around the micrometer sized aperture, consists in back-etching the insulating walls of the aperture while leaving the electrode material unaffected by the applied etching reactant. To this end, a SiN membrane is first coated on both sides with a carbon layer of a few tens of nanometers in thickness by electron beam evaporation (Fig.3.3 (a)). Undercoating with a thin adhesion layer of chromium has turned out to be useful. Subsequently an aperture of 1 µm in diameter is structured as shown in Fig. 3.3 (b) using a focused

---

[1] Silson Ltd, www.silson.com

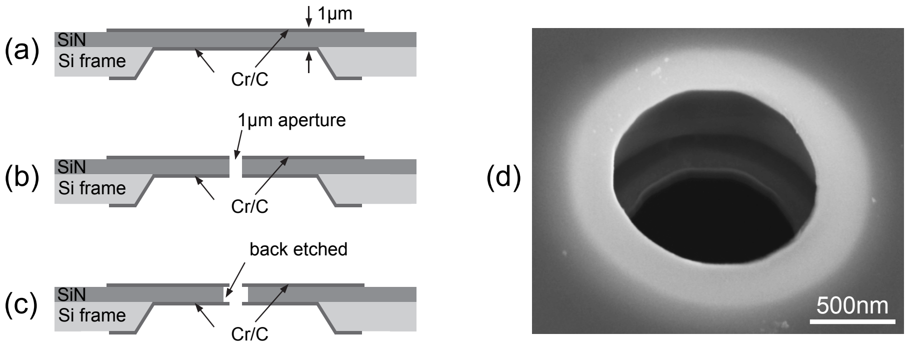


Figure 3.3: *Schematic drawings of the various steps for lens fabrication by BHF etching and an SEM image of the final lens structure. (a) Deposition of a carbon layer and a thin chromium undercoating on either side of a SiN membrane. (b) Ion-milling of 1 μm sized aperture. (c) Back-etching of the SiN walls around the aperture with BHF by a few hundred nanometers. (e) SEM image of the fabricated microlens, recorded at a tilt angle of 30°.*

gallium ion beam instrument (FIB). Since the undercutting of the SiN must occur in lateral direction, the applied etching process must be isotropic. Therefore we chose the method of chemical wet etching by buffered hydrofluoric acid (BHF) according to the reaction [23]:

$$\mathrm{Si_3N_4} + 18\,\mathrm{HF} \rightarrow 2(\mathrm{NH_4})_2\mathrm{SiF_6} + \mathrm{H_2SiF_6}\,. \tag{3.1}$$

The etch rate at room temperature is between 25 and 80 nm per hour [24, 25], thus requiring an etching time of several hours for back-etching the SiN by a few hundred nanometers. The chromium and carbon layers turned out not to be attacked by the etching process. For etching, the whole membrane is plunged in the BHF solution and agitated during the etching process to prevent sticking of contaminations to the membrane. The resulting structure is schematically depicted in Fig. 3.3 (c) and an SEM image of the resulting 1 μm sized aperture with the SiN walls back-etched by about 250 nm is shown in Fig. 3.3 (d).

This fabrication method is straightforward and could also easily be applied for the production of a larger number of lenses. However, the resulting structures sometimes suffer from contaminations around the rim of the aperture, leading to charging effects when implemented as a lens. The method can be improved with respect to contaminations by using another insulating material, e.g. silicon oxide, which is etched by hydrofluoric acid with higher rates.

3.2.2 Lens fabrication by deposition of a carbon flake

An alternative way to achieve that the electrodes around the apertures are free-standing is based on the idea of first removing the insulating material in a sufficiently large region and covering this only afterwards with an electrode layer into which the micrometer sized aperture is structured in a last step. This can in principle be realized in the following way: In a first step, a thin conductive layer is deposited on one side of the SiN membrane, just to enable imaging in the SEM and FIB (Fig. 3.4 (a)). A hole of several micrometers in size is milled through the membrane in a next step (Fig. 3.4 (b)). For the following process step, carbon is first evaporated on mica sheets. Subse-

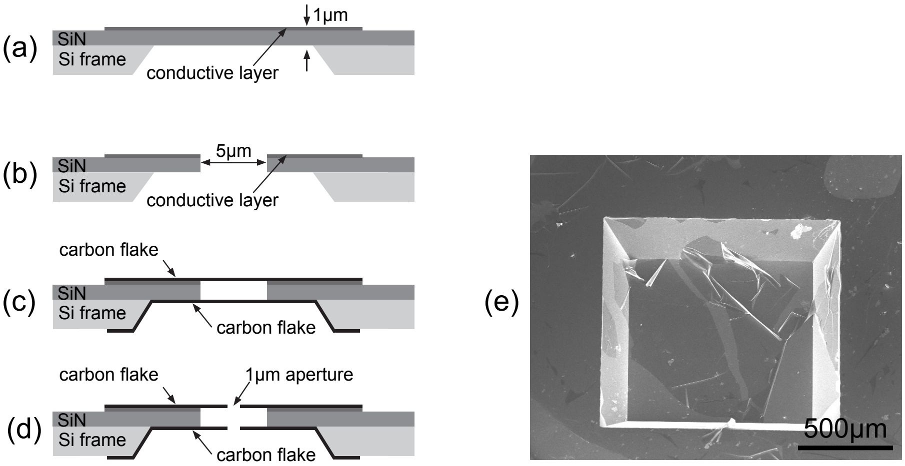


Figure 3.4: *Schematic drawings of the various steps for lens fabrication by covering with carbon flakes at both sides of the lens structure and an SEM image of the final lens structure. (a) Coating of the SiN membrane on one side with a conductive layer. (b) Structuring a circular 5 µm diameter aperture using a FIB. (c) Covering of the structure on both sides with a carbon flake. (d) Milling a 1 µm diameter aperture through the free-standing carbon layers using the FIB. (e) SEM image of the bottom side of the membrane covered with a carbon flake.*

quently, the carbon film is floated off onto a clean water surface [26] and can thus be deposited onto the upper side of the structure. The same is then done for the other side of the membrane (Fig. 3.4 (c)). As the final step, a hole of 1 µm diameter is ion-milled through the free-standing parts of both carbon films (Fig. 3.4 (d)). This method has been shown to work, yielding clean lens structures. However, since the SiN membrane on the lower side of the chip is recessed relative to the 200 µm thick silicon frame as can be seen in Fig. 3.2 (b), the deposition of the carbon flake on this side without crinkles and cracks turns out to be difficult (Fig. 3.4 (e).

Hence, this fabrication method has been varied, so that deposition of a carbon flake is necessary only on the upper side of the chip. For this method, the initial fabrication step consists in electron beam evaporation of a roughly 30 nm thick layer of amorphous carbon on either side of the SiN membrane (Fig. 3.5 (a)). Next, the FIB is used to remove the top carbon as well as the SiN layer within a circular region of 5 µm in diameter, whereas the carbon layer at the bottom side of the structure remains unchanged (Fig. 3.5 (b)). Subsequently, a carbon film is floated off a mica sheet and deposited onto the upper side of the structure (Fig. 3.5 (c)). Finally, a hole of 1 µm diameter is structured through the free-standing parts of both carbon films using the FIB (Fig. 3.5 (d)). An SEM image of such a final lens structure is shown in Fig. 3.5 (e). The lenses fabricated in this way show good performance in terms of stability and cleanliness.

3.2.3 Fabrication of asymmetric lenses

Another lens fabrication method that shall be described here does not involve back-etching or deposition of carbon flakes but consists merely in several steps of electron beam evaporation and ion milling. Charging of the insulating walls of the aperture is in this case prevented by making the aperture at the entrance side of the lens smaller than the aperture at the exit side. The first step

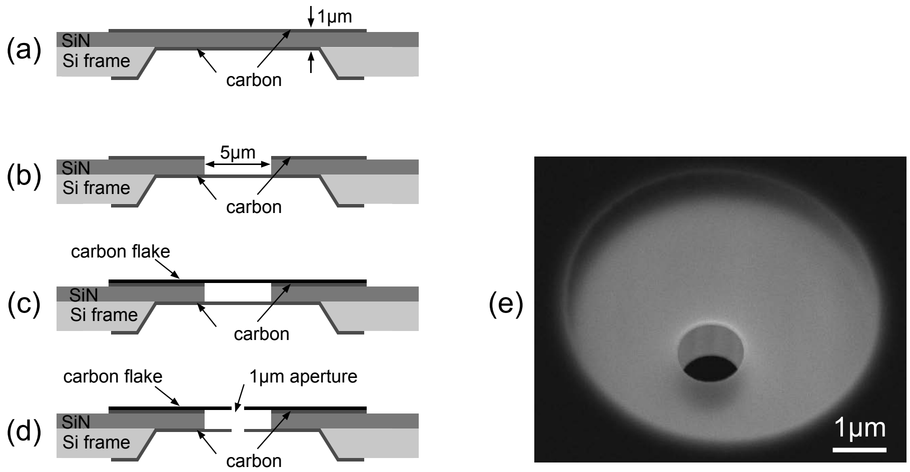


Figure 3.5: *Schematic drawings of the various steps for lens fabrication with a carbon flake and an SEM image of the final lens structure. (a) Carbon deposition on either side of a SiN membrane. (b) Removing the upper carbon and the SiN layer within a circular 5 µm diameter region using a FIB. (c) Covering of the structure with a carbon flake. (d) Milling a 1 µm diameter aperture through the free-standing carbon layers using the FIB. (e) SEM image of the fabricated microlens, recorded at a tilt angle of 30°. The penetration of the 12 keV electrons used in the SEM, allows recognizing the 5 µm diameter circular region where the carbon electrodes are free-standing. In this case, the lens aperture of 1 µm in diameter is not concentric with the 5 µm region but the distance to the SiN walls is large enough to prevent charging effects when implemented as a lens.*

consists again in evaporation of a few tens of nanometers of carbon on both sides of the silicon nitride membrane. In a next step, the upper carbon layer as well as the SiN are removed in a circular area of about 2 µm in diameter using the FIB. A smaller hole of 1 µm in diameter is then milled through the remaining carbon layer at the bottom side of the structure. This is done without changing the ion current and the sample position in the FIB instrument in order to ensure good alignment of the two apertures. Compared to the lenses with equally sized apertures, the voltage to be applied between the lens electrodes in order to focus a beam of electrons is higher for this asymmetric lens geometry. While this kind of microlens is useful for collimating an electron beam in the accelerating mode, charging effects may cause problems when the lens is operated in the decelerating mode, since in this case the electron trajectories are bent towards the insulating walls as it is shown in Fig. 3.1 (b). Another critical issue is to get the two apertures exactly concentric which is a crucial condition for good focusing properties.

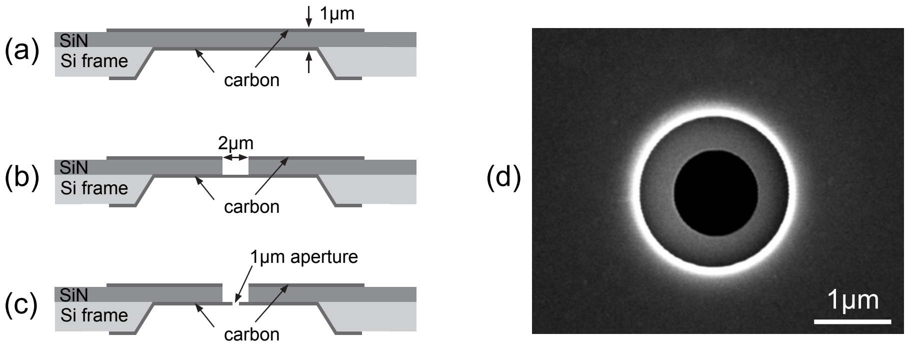


Figure 3.6: *Schematic drawings of the various steps for the fabrication of an asymmetric lens and an SEM image of the final lens structure. (a) Carbon deposition on either side of a SiN membrane. (b) Removing the upper carbon and the SiN layer within a circular 2 µm diameter region using a FIB. (c) Ion-milling a 1 µm diameter aperture through the carbon layer at the bottom side of the structure. (d) SEM image of the fabricated microlens.*

Chapter 4

Characterization of the microlens

4.1 Experimental implementation

Lenses fabricated as described above have been tested in the ultra-high vacuum system designed for experiments with coherent low-energy electrons that will be described in detail in Chapter 10. A W(111) field emission tip is used as a source for a divergent electron beam of high spatial and temporal coherence. Typical emitter currents are in the 10 to 200 nA range and the kinetic energy of the electrons at the lens entrance is well below 200 eV. The electron detector consists of a microchannel plate (MCP) followed by an electroluminescent layer on the vacuum side of a fibre optic plate (FOP). At the ambient pressure side of the FOP a CCD camera collects the emitted light. The detector resolution has been measured to be around 150 µm. A dedicated holder allowing for rapid vacuum transfer of microlenses fixes the position of the lens. The distance between lens and detector amounts to 75 mm. The electron source is mounted onto an x-y-z piezo-stage for precise alignment with the lens aperture. For the experiments described here, the source to lens distance varied between 5 and 30 µm, leading to kinetic energies of the electrons at the lens entrance between 60 and 150 eV.

4.2 Lens operation and overall performance

If the lens electrodes are both at ground potential and a negative voltage is applied to the emitter tip, a projection image of the second lens aperture is visible at the screen. Its magnification can be varied by changing the source-lens distance. Examples of such electron projection images are shown in Fig. 4.1 (a) and (e). Once a voltage is applied between the two lens electrodes, a focusing effect of the lens is observed. To ensure a field-free region beyond the lens, the second electrode is always kept at ground potential. The voltage applied at the first lens electrode is altered together with the voltage at the emitter tip to maintain the kinetic energy of the electrons at the lens entrance and the emission current constant. The polarity of the voltage at the first lens electrode can either be such that the electrons are decelerated when passing the lens (positive voltage), or such that they are accelerated (negative voltage). The effect of the lens is illustrated in Fig. 4.1 for both modes of operation. The images have been recorded using a lens fabricated by deposition of a carbon flake as shown in Fig. 3.5. While increasing the voltage between the two lens electrodes, one first observes that the projection image of the lens aperture at the screen decreases (Fig. 4.1 (b) and (f)). A further increase of the lens voltage reduces the divergence angle more and more until the electrons form a minimal spot at the detector (Fig. 4.1 (c) and (g)). With still higher lens voltages, a crossover located between lens and detector is achieved (Fig. 4.1 (d) and (h)). The distance between lens and crossover can even become smaller than the source-lens distance, resulting in an enlarged divergence angle of the beam as illustrated in Fig. 4.1 (d). Another focusing series for the decelerating mode of operation recorded with a lens fabricated by BHF etching as described in 3.2.1 is shown in Fig. 4.2. Due to small contaminations at the aperture edges, the image does not stay perfectly circular during focusing in this case. This focusing series also shows the outstanding importance of cleanliness of the aperture edges. Already slight contaminations, hardly recognizable in SEM, may cause distortions in the projection images during focusing, since the sensitivity of low-energy electrons to the buildup of electrical charge is much higher.

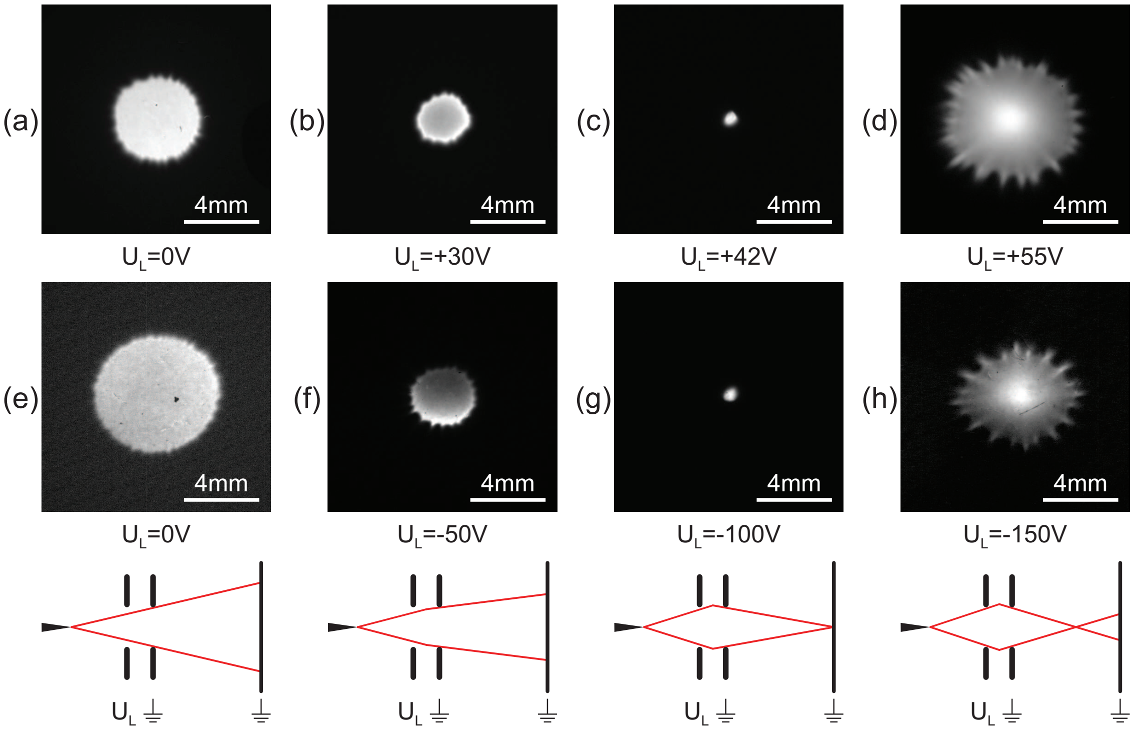


Figure 4.1: *Focusing series for the decelerating (top) and accelerating (middle) mode of a lens fabricated by carbon flake deposition with corresponding focusing situations (bottom). The distance between the virtual electron source and the first lens electrode amounts to 16 µm for the decelerating and 11 µm for the accelerating mode with kinetic energies of 93 respectively 95 eV at the lens entrance.*

4.3 Quantitative lens characterization

To evaluate whether downscaling the spherical aberrations by scaling down the size of the lens could actually be realized, the magnitude of spherical aberrations must be determined. Two quan-

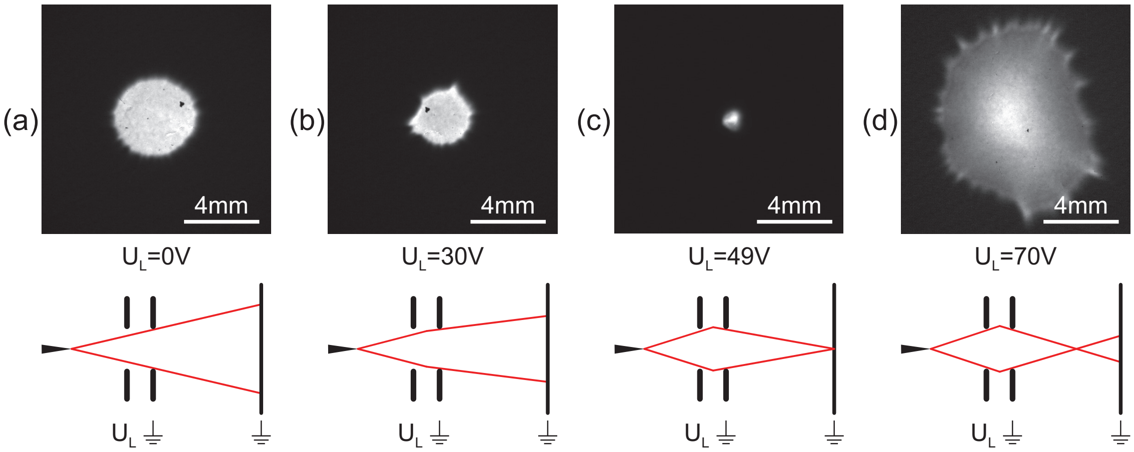


Figure 4.2: *Focusing series for the decelerating mode of a microlens fabricated by BHF etching (top) with corresponding focusing situations (bottom). The distance between the virtual electron source and the first lens electrode amounts to 17 μm with kinetic energies of 110 eV at the lens entrance.*

tities which are directly related to spherical aberrations have been measured. In a second step they were compared to the values obtained from ray tracing simulations, assuming an idealized lens deteriorated only by intrinsic spherical aberrations. The order of magnitude of the spherical aberration coefficients de facto realized in the experiments could thus be deduced.

4.3.1 Beam profile at a distant detector

An experimentally easily accessible quantity is the image size of the electron source in the plane of the electron detector which is the smallest achievable spot size at the detector. As the distance from the lens to the detector plane is large compared to the diameter of the lens aperture and to the source-lens distance, a minimal spot size at the detector is achieved when the beam leaves the lens almost parallel. Instead of determining spherical aberrations from the size of the spot obtained when a parallel incoming beam is focused, we did it the other way round. We instead measured the minimal achievable divergence angle when collimating a beam emitted by an almost perfect point source. Thus, the minimal spot size at the detector is directly related to the spherical aberrations of the lens.

The minimal achievable spot size was measured for the decelerating mode of the lens. Each measurement started with taking a projection image of the lens aperture. This allows determining the distance between the virtual electron source and the first lens aperture [27]. Next, the lens voltage was adjusted for obtaining a minimal spot at the distant detector. In order to get the size of the minimal spot and the projection image, the contour line corresponding to half the maximum intensity value in the image was determined. An ellipse was then fitted to the contour line and the spot diameter was taken as the mean of its minor and major axis.

It turns out, that the average spot size at the detector varies from 0.5 to 1.1 mm measured with a detector resolution of 120 µm. Thus, assuming a Gaussian spot profile the broadening of the spot imposed by the finite detector resolution amounts to less than 3% and will not be accounted for in the following. In addition to spherical aberrations there are several other intrinsic effects contributing to the size of the spot visible at the detector, namely chromatic aberrations, diffraction of the beam at the lens aperture and finite source size. For an estimation of chromatic aberrations the energy spread of 0.1% of the electrons [28] must be taken into account as well as the stability of the lens voltage. Ray tracing simulations show, that this leads to an enlargement of the spot size at the detector by about 15 μm. We estimated the influence of diffraction for the decelerating mode of the lens by assuming a parallel beam being diffracted at the second aperture of the lens. This leads to a width of the central maximum of the diffraction pattern of 15 μm for 50 eV electrons. The magnified image of the virtual electron source contributes with less than 10 μm to the spot size, as the size of the virtual electron source is well below 1 nm; in fact it has been measured to be of atomic dimension [29]. We can thus conclude that chromatic aberrations, diffraction at the lens aperture and finite source size are negligible contributions to the minimal spot size at the detector.

4.3.2 Beam profile at 200 μm distance beyond the lens

A more direct way to determine aberrations is to measure the beam diameter in a plane much closer to the lens than the 75 mm distant detector plane. This has been done by scanning a sharp edge perpendicular to the optical axis through the beam. The beam diameter in the plane of the edge is given by the displacement of the edge from the position where all electrons are blocked to the position where all electrons pass the edge. We fabricated such sharp edge by ion-milling a rectangular window into a 20 nm thick carbon foil, opaque for low-energy electrons. This was mounted onto the movable sample piezo-stage and the window was positioned into the electron beam 200 μm beyond the lens. Precise motion of the edge perpendicular to the optical axis is realized with a piezo-scanner exhibiting a scan-range of 4 μm. While moving the sample with a scan speed of typically 800nm/s, the total intensity at the detector was recorded. The lens voltage was then adjusted for the steepest slope in the intensity versus edge position profile while repeatedly scanning through the beam.

4.3.3 Ray tracing simulations

Calculation of electrostatic fields and electron trajectories have been done using the software package SIMION 8.0[1] to solve the Laplace equation using the finite-difference method. The lens was modelled as two planar electrodes separated by 1 μm and exhibiting concentric apertures of 1 μm in diameter with field-free regions on either side of the lens. Cylindrical symmetry reduces the computational effort since the Laplace equation must only be solved in two dimensions. The electrical field distributions were calculated on a 15000 × 1000 array with a grid size of 5 nm. Ray tracing was done with a fourth order Runge-Kutta algorithm implemented in SIMION.

---

[1] SIMION 8.0, Scientific Instrument Services, Inc., 2003-2006

Calculation of spherical aberration coefficients

Spherical aberrations cause paraxial and marginal rays to be focused at different positions, as illustrated in Fig. 4.3. The distance between the paraxial focus J and the focus J_n of a ray passing the lens at a larger distance from the optical axis is called longitudinal spherical aberration ΔJ_n of the ray, see Fig. 4.3 (a) for denotation. In the case of a parallel incoming beam, the longitudinal spherical aberration of a ray ΔJ_n can be expressed as a power series of the ray's semi angular aperture θ_n (see also Fig. 4.3 (b)). As ΔJ_n is a symmetric function of θ_n, the odd power terms vanish, which leads to:

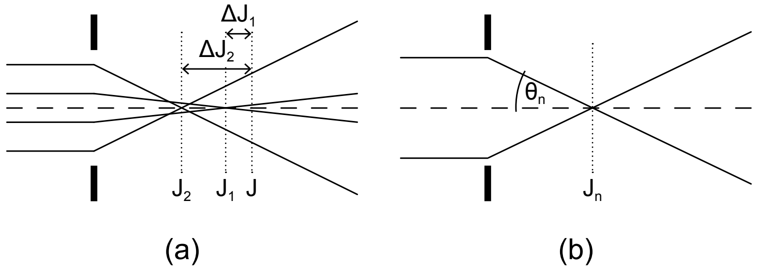


Figure 4.3: (a) Schematic drawing showing two rays with different initial distances from the optical axis which are focused at different positions J_1 and J_2. The paraxial focus is denoted as J. ΔJ_1 and ΔJ_2 are called the longitudinal spherical aberration of the two rays. (b) The position of the focus J_n depends on the semi angular aperture θ_n of the ray.

$$\Delta J_n = C_s\theta_n^2 + c_4\theta_n^4 + c_6\theta_n^6 + ..., \tag{4.1}$$

with the coefficient C_s being the primary spherical aberration coefficient [30]. The latter was calculated by combining ray tracing simulations with a least square fit, similar to methods described elsewhere [31]. Calculations were carried out for a parallel incident ray, both for the accelerating as well as for the decelerating mode of the lens. For this, a parallel beam of 110 electrons has been generated, such that the lens aperture was completely filled. For each particle, the focus position J_n and the elevation angle θ_n were computed. The position of the paraxial focus J was approxi-

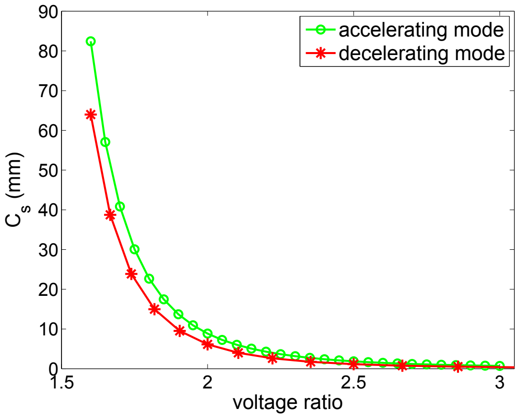


Figure 4.4: Primary aberration coefficient C_s as a function of the voltage ratio for the decelerating and accelerating mode of the lens.

mated by the focus of the ray with an initial distance to the optical axis of 5 nm. In this way, the longitudinal aberration $\Delta J_n = J - J_n$ was determined for each particle. A polynomial of sixth order was then fitted to the values $\Delta J_n(\theta_n)$, yielding the primary aberration coefficient C_s. The described procedure was applied to a series of lens voltages for the decelerating and accelerating mode respectively. In the following, voltages are always taken relative to the potential where the kinetic energy of the electrons is zero. Thus, the lens is completely described by its geometry and the voltages ratio between the two electrodes [30]. The dependence of C_s on the voltage ratio for the accelerating and decelerating mode are shown in Fig. 4.4 for voltage ratios related to realistic experimentally accessible values. The aberration coefficients C_s are in the range of millimeters for both modes and decrease with increasing focusing strength of the lens. Here, C_s denotes the aberration coefficient referred to the image side. Thus, the value for the lateral size of the focal spot in the plane of least confusion d_s can be calculated using the relation $d_s = 0.5 C_s \theta^3$, where θ is the angular aperture at the image side of the lens [32]. For a crossover several ten micrometers away from the lens, this yields a spot size of several ten nanometers.

Calculation of the spot size at a given distance

For comparison with experimental values the size of the image of the electron source was calculated in the detector plane 75 mm beyond the lens and in the plane located 200 µm beyond the lens. As the main contribution to the finite spot size in a plane distant to the lens is due to spherical aberrations, other factors influencing the spot size, as there are chromatic aberrations, finite source size and diffraction effects, were neglected in the simulations. A monochromatic point source was thus assumed. Since typical experimental tip to lens distances are larger than 5 µm, the electrostatic field near the first lens-aperture and the strong field close to the field emission tip do not influence each other, so that they can be treated as two separated electron-optical components. The electron trajectories in the immediate vicinity of the field emission tip are slightly curved as most of the potential drops in front of the tip. However, further away from the tip, at the position of the lens, the trajectories are straight lines. At the lens entrance they seem to emanate from a virtual source located a short distance behind the physical tip apex as will be discussed in more detail in 10.2.2. There is thus no need to include the field emission tip in the electrode array. Instead, the region at the entrance side of the lens can be assumed as field-free. Electrons originate at a point corresponding to the position of the virtual source and propagate straight with given kinetic energies and divergence angles towards the lens. As the region behind the lens is field-free, the beam diameter in the planes of interest up to the 75 mm distant detector was calculated by extrapolating the trajectories. In accordance with the experimental conditions, a sufficiently large divergence angle was selected to ensure that the electrons completely fill the lens aperture. The distance between subsequent electrons in the plane of the lens aperture was set as close as possible, that is to one grid unit corresponding to 5 nm. For several distances between electron source and first lens-aperture the spot size in the plane under consideration was calculated for a set of various lens voltages. In this way, the minimal possible spot size was computed as a function of source-lens distance for both planes evaluated in the experiments, which is at 75 mm and 200 µm beyond the microlens.

4.4 Results and discussion

4.4.1 Minimal spot size at the 75 mm distant detector

Measured spot sizes at the detector as a function of source-lens distance are depicted in Fig. 4.5 together with the corresponding simulated values. Simulations predict a minimal spot size at the detector between 0.2 and 1.1 mm in diameter at the detector for source-lens distances ranging from 5 to 30 µm. As expected, the increase of the divergence angle with decreasing source-lens distance leads to a larger spot size. The simulated values exhibit some fluctuations which are due to the limited precision in the determination of the lens-voltage required to form a minimal spot. Two series of experimental measurements of the spot size at the detector are also plotted in Fig. 4.5. For series 1 several measurements have been carried out with one and the same microlens, fabricated in the manner described in 3.2.2 by deposition of a carbon flake. In contrast to this, the eight data points of series 2 have been obtained from measurements with eight different microlenses. Some of them have been produced by deposition of a carbon flake as well, others have been produced by BHF etching as described in 3.2.1; however, all of them exhibit the same geometry. The measured values for the minimal spot size at the detector show the predicted dependency. The experimental data points however are shifted by 0.3 mm towards higher values compared to the simulated ones. The measured spot sizes assume values between 0.5 and 1.1 mm corresponding to a divergence angle of the collimated beam between just 6 and 14 mrad. The experimental values of series 2 obtained with different lenses exhibit similar characteristics. This implies that microlenses with comparable quality can now routinely be fabricated in a reproducible way. Moreover, as different lens fabrication methods result in similar lens performances, they appear to be equivalent with respect to the focusing properties of the lenses.

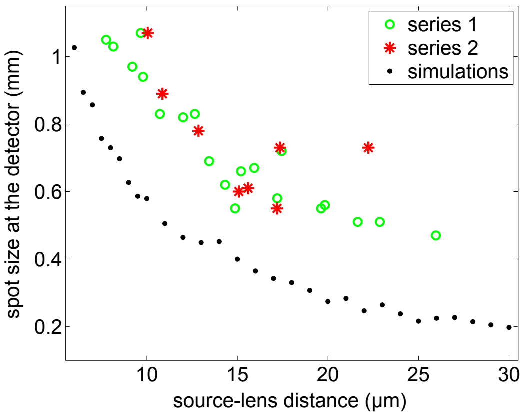


Figure 4.5: *Comparison between simulated and experimental values for the minimal spot size at the detector. Data points of series 1 have been obtained from measurements with one individual lens while the data for series 2 correspond to measurements with eight different lenses.*

4.4.2 Minimal spot size in a plane 200 µm beyond the lens

The simulated spot-sizes in the plane 200 µm beyond the lens vary from 3 µm diameter for a source-lens distance of 6 µm down to 0.6 µm diameter for a source-lens distance of 30 µm. Experimental values have been obtained as described above (see also Fig. 4.6 (a)). For a source-lens distance of 14 µm and 92 eV electrons at the lens entrance, a lens-voltage of 48 V was found to generate a minimal spot. In Fig. 4.6 (b) the total intensity at the screen, averaged over all pixels of the CCD chip, is plotted versus the position of the edge moved perpendicular to the beam with a scan speed of 800 nm/s. From the intensity profile a beam diameter of 1.8 µm has been derived. Corresponding simulations assuming identical conditions reveal a value of 1.2 µm. As the distance of 200 µm is still large compared to the 1 µm diameter of the lens apertures, focusing of the beam towards this plane is impossible. Instead, the beam diameter is minimal when the beam leaves the lens as parallel as possible. Thus, the focusing conditions for a minimal beam diameter are just as in the case of a 75 mm distant detector as described and discussed above. In fact, the values for the beam divergence angle obtained by the two methods are in very good agreement; they amount to 9 mrad for a source-lens distance of 14 µm.

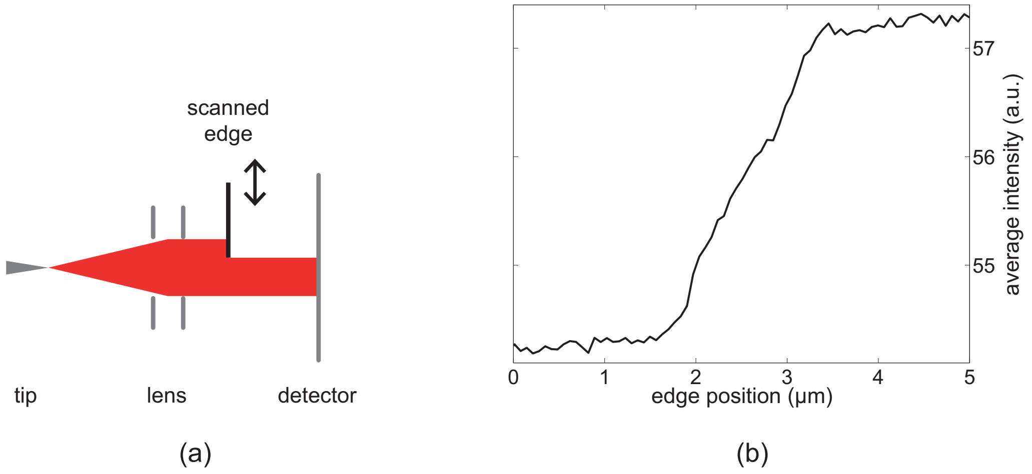


Figure 4.6: (a) To measure the beam diameter 200 µm beyond the lens, a micro-machined edge is scanned through the beam while the total intensity at the screen is recorded. (b) Measured intensity at the screen as a function of edge position. The source-lens distance amounts to 14 µm. A beam diameter of 1.8 µm is deduced from the width of the step profile.

Chapter 5

Conclusions and outlook

It could be shown that electrostatic lenses of micrometer dimensions can readily be fabricated and exhibit reproducible performance. Comparisons with ray tracing simulations show that the concept of scaling down spherical aberrations by decreasing lens dimensions can in fact be realized. Effects like misalignment of source and lens, residual vibrations and ac-magnetic fields, deviations from perfect lens-symmetry and contaminations of the apertures have not been considered at all in our simulations. Therefore, the experimental values for the spot size are somewhat larger than the values predicted by simulations. Although the aforementioned effects do reduce lens performance, their adverse effect turns out to be small compared to the benefit in terms of spherical aberrations due to down-scaling lens dimensions. From a comparison to ray tracing simulations one thus can conclude that a simple micrometer-sized lens exhibits aberration coefficients in the range of several millimeters. With this, spot sizes in the nanometer range are expected to be generated at distances of several tens of micrometers beyond the lens exit. Direct measurement of the focal spot is possible by scanning an edge with nanometer precision across the beam and measuring the total intensity at the detector as it has been described in 4.3.2. This method is only limited by positioning issues, occurring when two planar structures are approached towards each other down to distances in the micrometer range.

While the mere size of the aberration coefficient and the thus achievable size of the focal spot is just one aspect, an even more important feature of the scaling concept is that the electron beam can always be kept close to the optical axis. Using the microlens to collimate the beam, it never spreads out to more than 1 µm from the optical axis up to a distance of 200 µm beyond the lens exit. This is advantageous if the microlens is implemented as one component of an electron optical systems. If an objective lens is positioned beyond the microlens, it perceives a paraxial beam with a broadening of less than 10 mrad. No beam limiting aperture to mitigate aberrations of the objective lens but decreasing brightness at the same time would consequently be needed. As it has been shown by Chang [18], an electron column with aberration coefficients in the range of conventional magnetic columns of comparable working distances can be achieved by combining a low-aberration microlens with a simple electrostatic einzel lens, exhibiting dimensions in the range of 10 to 100 µm. Such an electron column would have an overall size of about one centimeter, thus offering the possibility to build a miniaturized low-energy electron scanning microscope. Reduced penetration depth and radiation damage as well as less charging effects due to the low-energy of the electrons such a device would have a number of interesting applications. A focused beam of low-energy electrons should be useful for surface science and examination of organic and insulating materials. Moreover, one can also think of lithographic applications, where proximity effects, limiting the resolution in present technologies could be significantly reduced.

Further improvement of the microlens itself can be achieved by adjusting the several geometrical parameters of the lens. Optimizing thickness of the electrodes, the aperture sizes and the thickness of the insulating layer, spherical aberrations can be reduced further. Moreover, fabrication of a three-electrode microlens appears feasible by the methods established here. Apart from being advantageous with respect to lens aberrations, a three-electrode lens, when operated as an immersion lens, offers an additional degree of freedom concerning the final kinetic energy of the electrons at the lens exit. When implemented in the low-energy electron diffraction experiment the kinetic energy and with this the associated wavelength could be adjusted within the lens instead of post-accelerating the electrons towards the sample (compare 10.3.2).

The fabricated micrometer sized two-electrode lenses with a performance as shown here constitute a pivotal component of the coherent low-energy electron diffraction experiment that has been set up within the course of this work. Their focusing properties allow collimating the beam of low-energy electrons to a final divergence angle of several milliradians only, small enough to fulfil the requirements for a coherent diffraction experiment. By downscaling both, the source as well as the electron lens, a spherical wave of high spatial coherence as it is emitted from the field emitter with an effective source size of near atomic dimension, can be transformed by the microlens into a plane wave preserving the high spatial coherence of the electron wave.

Part II

Diffraction of coherent low-energy electrons on individual molecules

Chapter 6

Principle of coherent diffraction imaging

Coherent diffraction microscopy has been developed as a technique for imaging non-crystalline specimens during the past ten years. It consists in reconstructing the image of a specimen from its far-field diffraction pattern by means of numerical algorithms. In case of a plane incident wave, the diffracted wave in the far-field corresponds to the Fourier transform of the wavefield in the plane just at the exit of the object [33]. Since this object exit wave contains the information about the interaction of the incident wave with the object it can be considered as an image of the latter. Unfortunately, when recording the diffraction pattern, only the squared amplitudes of the Fourier transform of the object exit wave can be measured, while the associated phases are lost. As a consequence, direct recovery of the object exit wave by inverse Fourier transformation is not possible. This is the well known *phase problem* and will be discussed in detail in Chapter 8.

However, the diffraction pattern of a non-periodic specimen is continuous which, under certain conditions, allows the phases to be retrieved numerically by iterative algorithms. The idea can be traced back to Sayre [34] and algorithms for the reconstruction of an object from the magnitude of its Fourier transform have been developed during the 1970s [35, 36]. Yet, the incident radiation must exhibit high spatial and temporal coherence, as will be expounded in Chapter 9. Thus, only with the development of synchrotron sources, providing highly coherent x-ray radiation and in addition high beam brightness, the technique of coherent diffraction imaging (CDI) was realized experimentally. The first successful reconstruction of the image of a two-dimensional object from its x-ray diffraction pattern could be achieved in 1999 by Miao and co-workers [37] who used synchrotron radiation with a wavelength of $\lambda = 1.7\,\text{nm}$ and 100 nm sized gold dots arranged on a silicon-nitride membrane as a test sample. Other applications of CDI using x-rays of different wavelengths have been reported thereafter, yielding images of fabricated micro-structures [38], of nanocrystals and nanoclusters [39, 40, 41] and also of biological samples like bacteria and whole cells [42, 43].

Extension of the method to three-dimensional imaging by recording diffraction patterns of the sample through a set of rotation angles has also been proven to be possible [44]. Apart from experiments with x-rays, coherent diffraction microscopy has also been performed using high-energy electrons. In 2003 Zuo et al. succeeded in imaging a carbon nanotube with atomic resolution using 200 keV electrons [45] and other publications followed [46, 47]. Besides, the experiments for imaging the atomic structure of biological molecules, which are planned at the XFELs, are also based on the technique of coherent diffraction imaging [7, 44, 48, 49].

Chapter 7

Some aspects of scattering theory

For the description of the interaction of electrons with matter one distinguishes between elastic scattering processes, where the initial kinetic energy of the particle is conserved and inelastic scattering processes, involving a change in kinetic energy. While elastically scattered electrons are the ones contributing to the signal in coherent diffraction imaging, the inelastically scattered electrons only give raise to an incoherent diffuse background signal and, moreover, may also damage the sample due to energy transfer. For electrons with high energies as conventional for TEM, i.e. 100 - 200 keV, as well as for x-rays, the inelastic scattering processes prevail, and by the time sufficient elastic scattering events have occurred to produce data with a good signal-to-noise ratio, a biological sample will be completely destroyed [4, 8]. Fortunately, things appear to be more favourable for low-energy electrons, with energies in the range of 50 - 250 eV as it will be discussed in some detail in the following section. A concept based on quantum mechanical scattering theory to explain the formation of a diffraction pattern and its interpretation will be provided in 7.2.

7.1 Interaction of low-energy electrons with matter

Elastic scattering of high-energy electrons occurs by Coulomb interaction between the electron and the electrostatic potential of the nuclei of the sample atoms. Screening of the nucleus Coulomb potential by the electrons of the sample atoms can be neglected for electron energies above 100 keV resulting in a Rutherford scattering cross section [50]. For lower electron energies, this approximation becomes invalid and screening must be taken into account. This is usually done by modifying the Coulomb potential introducing an exponential screening factor [51]:

$$V(r) = -\frac{e^2 Z}{4\pi\epsilon_0 r} e^{-r/R} , \tag{7.1}$$

where $R = a_H Z^{-1/3}$ (a_H being the Bohr radius) denotes the screening radius of the atom. Calculations of the differential cross-sections for a screened Coulomb potential using the Pauli-Dirac equation result in the so-called Mott cross sections. They provide a description for the elastic scattering of electrons by a solid, predicting an increasing rate of elastic scattering events with decreasing electron energy [52]. For energies below 50 eV, however, the Mott cross sections overestimate the scattering rate [50] and other models have been proposed for a description of scattering of these very low-energy electrons at solids [52], involving quasi-elastic scattering on acoustic phonons. Therefore, these models are not applicable to scattering on individual molecules.

Inelastic scattering, and hence energy transfer from the incident electron to the object, may occur via several different processes. When a bulk solid is considered, the jellium model thus being applicable as an approximation, these include electron-electron collisions between the incident electron and a free electron of the atomic jellium as well as plasmon excitation. In the case of a single molecule however, the energy transfer can only occur by excitation of electrons of the sample atoms to higher states or by even removing an electron from an atom resulting in ionisation. By comparing the calculated cross sections for these inelastic events with the cross sections predicted for elastic scattering, one can conclude, that the ratio between the two is more favourable for electron energies in the range of 100 eV than for higher energies [52].

Although the theoretical models sketched here, describing the interaction of an electron with solid matter, have proven to be very useful for predicting interactions of high-energy electrons with bulk solid state samples, they may fail for low-energy electrons interacting with individual molecules prepared such that they are suspended over holes in support films. This is also indicated by experiments with very low-energy electrons (energies of 1 - 20 eV), showing that inelastic scattering from a DNA molecule may occur by other mechanisms than ionisation, causing strand breaks and consequently leading to destruction of the molecule [53]. Moreover, since these mechanisms are based on resonance effects, they strongly depend on the molecular environment. As secondary electrons have energies just in the range between 1 and 20 eV, these processes can also contribute to radiation damage in the case of higher primary electron energies.

Experimental evidence for non-destructive imaging of individual DNA molecules with low-energy electrons, i.e. electrons with energies between 50 end 250 eV, could recently been provided [9]. DNA molecules suspended over holes in a carbon film could be exposed to low-energy electron radiation in a LEEPS microscope [10] for hours without any detectable changes being induced. The applied electron current and illumination time corresponded to 10^8electrons/ nm^2, which is a factor 10^5 larger than the maximal dose applicable to a biological sample using x-rays or high-energy electrons [4, 8]. Preparation of the sample molecules over holes reduces the emission of secondary electrons from the sample support film and thus prevents indirect radiation damage by effects as mentioned above. The probability for inelastic scattering of low-energy electrons at the molecule itself is obviously small compared to the rate of elastic scattering, hence allowing the collection of good signal-to-noise data without destroying the specimen.

7.2 The Born approximation

In order to explain the relation between the object and its far-field diffraction pattern, a brief survey of a quantum mechanical description of the scattering process shall be given here. This shall also clarify, how the image obtained by application of the reconstruction algorithms to a diffraction pattern must be interpreted.

The interaction of an electron wave $\Psi(\mathbf{r},\mathbf{t})$ with the electrostatic potential $\varphi(\mathbf{r})$ of a specimen is described by the Schrödinger equation:

$$-\frac{\hbar^2}{2m}\,\nabla^2\,\Psi(\mathbf{r},t) + e\varphi(\mathbf{r})\Psi(\mathbf{r},t) = -\mathrm{i}\hbar\frac{\partial\Psi(\mathbf{r},t)}{\partial t}\ . \tag{7.2}$$

As long as only elastic scattering is considered, the time independent form of the Schrödinger equation holds:

$$\nabla^2\Psi(\mathbf{r}) + \frac{2me}{\hbar^2}\,[E+\varphi(\mathbf{r})]\,\Psi(\mathbf{r}) = 0\,, \tag{7.3}$$

where eE is the kinetic energy of the electron in the field-free space, i.e. for $\varphi = 0$. The time independent Schrödinger equation (7.3) has the form of a wave equation

$$\left(\nabla^2 + k^2\right)\Psi(\mathbf{r}) = 0\,, \tag{7.4}$$

where the wave vector is given by

$$k^2 = \frac{2me}{\hbar^2}\,(E+\varphi\,(\mathbf{r}))\ . \tag{7.5}$$

Thus, in the case of vanishing potential $\varphi(\mathbf{r}) = 0$, a possible solution of the time independent Schrödinger equation is a planar wave

$$\Psi(\mathbf{r}) = \Psi_0\ \exp\left(\mathrm{i}\mathbf{kr}\right)\ . \tag{7.6}$$

Another solution is a spherical wave emanating from a point taken to be the origin:

$$\Psi(r) = \Psi_0\,\frac{\exp\left(\mathrm{i}kr\right)}{r}\,, \tag{7.7}$$

with the wave vector $\mathbf{k}$ always pointing in radial direction. In order to solve the wave equation for the case of a non-vanishing potential $\varphi(\mathbf{r})$, i.e. scattering occurs, (7.3) can be written in the corresponding integral form by means of the Green's function $G(\mathbf{r},\mathbf{r}')$:

$$\begin{aligned}\Psi(\mathbf{r}) &= \Psi^{(0)}(\mathbf{r}) + \Psi^{(1)}(\mathbf{r})\\ &= \Psi^{(0)}(\mathbf{r}) + \int G(\mathbf{r},\mathbf{r}')\,\varphi(\mathbf{r}')\,\Psi(\mathbf{r}')\,d\mathbf{r}'\,,\end{aligned} \tag{7.8}$$

where $\Psi^{(0)}(\mathbf{r})$ denotes the plane or spherical incident wave. The appropriate form of the Green's function in this case is a spherical wave emanating from point $\mathbf{r}'$

$$G(\mathbf{r},\mathbf{r}') = \frac{\exp\left(\mathrm{i}k|\mathbf{r}-\mathbf{r}'|\right)}{4\pi|\mathbf{r}-\mathbf{r}'|}\ . \tag{7.9}$$

The integral equation (7.8) can in general not be solved analytically since the wave function $\Psi(\mathbf{r})$ appears both, on the left hand side as well as within the integral. In order to obtain a first approximation, it can be assumed that the amplitude of the scattered wave is much smaller than the incident wave amplitude and that $\Psi(\mathbf{r})$ within the integral may be replaced by the incident wave $\Psi^0(\mathbf{r})$, yielding

$$\begin{aligned}\Psi(\mathbf{r}) &= \Psi^{(0)}(\mathbf{r}) + \Psi^{(1)}(\mathbf{r}) && (7.10)\\ &= \Psi^0(\mathbf{r}) + \int G(\mathbf{r},\mathbf{r}')\,\varphi(\mathbf{r}')\,\Psi^0(\mathbf{r}')\,d\mathbf{r}' \,. && (7.11)\end{aligned}$$

This is known as the *first Born approximation.* Higher order terms of the wave function may be calculated by a recurrent scheme; the second approximation is given by replacing $\Psi(\mathbf{r})$ in the integral by $\Psi^0(\mathbf{r}) + \Psi^1(\mathbf{r})$ and so on. For a plane incident wave with wave vector $\mathbf{k_0}$, the First Born approximation gives

$$\Psi(\mathbf{r}) = \exp(\mathrm{i}\mathbf{k_0}\cdot\mathbf{r}) + \frac{1}{4\pi}\int \frac{\exp(ik|\mathbf{r}-\mathbf{r}'|)}{|\mathbf{r}-\mathbf{r}'|}\,\varphi(\mathbf{r}')\exp(\mathrm{i}\mathbf{k_0}\cdot\mathbf{r}')\,d\mathbf{r}' \,. \tag{7.12}$$

If the point of observation $\mathbf{r}$ is at a distance R large compared to the dimensions of the object, the expression for the scattered wave can further be simplified to:

$$\Psi^{(1)}(\mathbf{r}) = \frac{1}{4\pi}\frac{\exp(\mathrm{i}kR)}{R}\int \varphi(\mathbf{r}')\exp(-\mathrm{i}\,\mathbf{q}(\mathbf{r})\cdot\mathbf{r}')\,d\mathbf{r}' \,, \tag{7.13}$$

denoting with $\mathbf{q} = \mathbf{k} - \mathbf{k_0}$ the difference between the scattered and the incident wave vector. The former points in the direction of the point of observation and, since $|\mathbf{k}| = |\mathbf{k_0}|$ for elastic scattering, it is uniquely determined by $\mathbf{r}$. The relation between the vector $\mathbf{q}$ and the direction of observation is given by

$$\mathbf{q} = \frac{k_0}{r}\mathbf{r} - \mathbf{k_0} \,. \tag{7.14}$$

Equation (7.13) shows the important result that, in the case of a plane incident wave and applying the first Born approximation, the scattered wave $\Psi^{(1)}(\mathbf{r})$ at a large distance from the object conceived as a function of the scattering vector $\mathbf{q}$ corresponds to the Fourier transform of the scattering potential. It is quite common to refer to the scattered wave not with respect to the coordinates $\mathbf{r}$ in real space but to use the coordinates $\mathbf{q}$ in reciprocal space instead and hence defining a scattering amplitude $F(\mathbf{q})$ which contains the characteristic scattering from the object's potential field:

$$F(\mathbf{q}) = \frac{1}{4\pi}\int \varphi(\mathbf{r}')\exp(-\mathrm{i}\mathbf{q}\cdot\mathbf{r}')\,d\mathbf{r}' \,. \tag{7.15}$$

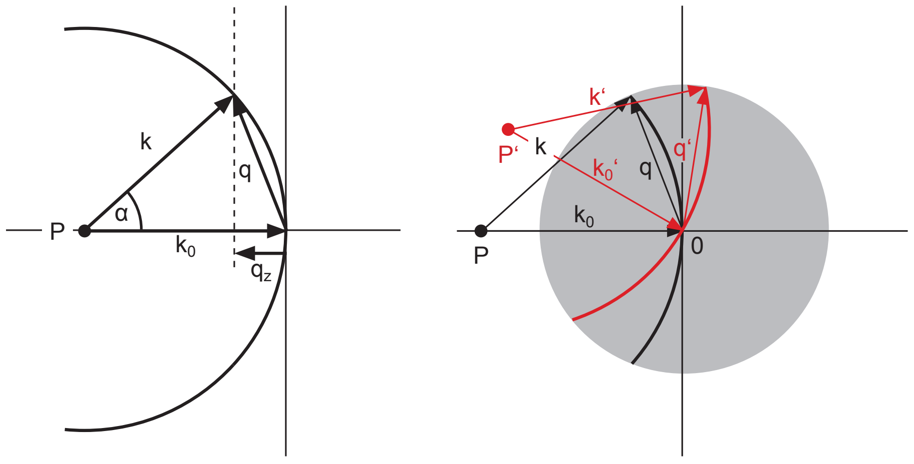
The first Born approximation comprises the assumption that the amplitude of the scattered wave is so small, that only scattering of the incident wave itself must be considered. Thus, the first Born approximation is a single-scattering approximation.

7.2.1 The Ewald sphere

Since we consider the case of elastic scattering, momentum conservation holds and it is

$$|\mathbf{k}| = |\mathbf{k_0}| \,. \tag{7.16}$$

Thus, the vectors $\mathbf{q} = \mathbf{k} - \mathbf{k_0}$, giving the difference between scattered and incident wave vector, can be represented in reciprocal space as it is illustrated in Fig. 7.1 (a). The vectors $\mathbf{q}$ are all pointing from the origin to the surface of the so called Ewald sphere centred at a point P and having a radius $|\mathbf{k_0}|$. The point P is determined by the incident wave vector; its distance from the origin is $|\mathbf{k_0}|$. Thus, according to (7.13), by measuring the intensities of a far-field diffraction pattern for a given direction of the incident wave, one obtains the values of the three-dimensional Fourier transform of the scattering potential at points lying on a spherical surface in reciprocal space. By rotating the sample about an axis perpendicular to the direction of the incident wave vector or, alternatively, by changing the direction of the incident beam, the sphere is rotated about the origin in reciprocal space as depicted in Fig. 7.1 (b), hence making more values of the function $F(\mathbf{q})$ accessible. Accordingly, a three dimensional data set can be recorded, containing information about the Fourier intensities of the scattering potential up to a maximal q-value, which depends on the length of the wave vector $\mathbf{k}$ and the detector geometry.



(a) (b)

Figure 7.1: *(a) Construction of the Ewald sphere. Since the length of the wavevector is not changed by elastic scattering, all the scattering vectors* $\mathbf{q} = \mathbf{k} - \mathbf{k_0}$ *point to a spherical surface centred at point P with radius* k_0*. The component of the scattering vector in z-direction can not be neglected for large scattering angles* α*. (b) The data obtained by recording the diffraction pattern up to a maximal scattering angle* $\mathbf{q}$ *correspond to values of the scattering amplitude lying on the Ewald sphere. By altering the direction of the incident wave vector, the Ewald sphere is rotated about the origin making another part of the reciprocal space accessible for data collection.*

If only scattering through small angles is recorded, as it is usually the case in high-energy electron diffraction experiments, the curvature of the Ewald sphere across the observed region may be neglected and it can be approximated by a plane. The recorded diffraction pattern can thus be considered as a planar section through the origin of the three dimensional Fourier transform of the scattering potential [54]. Choosing the direction of the incident wave as z-direction of the coordinate system, and denoting the coordinates in reciprocal space by q_x, q_y and q_z, this can be written as

$$F(q_\mathrm{x}, q_\mathrm{y}, 0) = \iint \exp\left(-i(q_\mathrm{x} x + q_\mathrm{y} y)\right) \cdot \int \varphi(x, y, z)\, dz\, dx\, dy\,. \tag{7.17}$$

This is the two-dimensional Fourier transform of the projection of the potential $\varphi(\mathbf{r})$ along the z-axis. However, in order to obtain a resolution on the atomic level with low-energy electrons, i.e. with a radiation of wavelengths between 1 and 2 Å, the detection of diffraction signal through considerably large angles is necessary. Therefore, the curvature of the Ewald sphere can not be neglected and the measured diffraction intensities do not correspond to a planar section of the reciprocal space. As can be inferred from the illustration in Fig. 7.1, the maximal distance of the sphere from the plane $q_z = 0$ is determined by the maximal diffraction angle α_max to:

$$q_\mathrm{z,\,max} = k\,(1 - \cos(\alpha_\mathrm{max}))\ , \tag{7.18}$$

which deviates considerably from zero for maximal diffraction angles of $\alpha_\mathrm{max} = 25°$, as they are required for atomic resolution.

For a two-dimensional object, it follows from (7.13), that the three-dimensional Fourier transform is independent of q_z. Thus, in this case, it does not make a difference if the intensities are measured at $q_\mathrm{z} = 0$ or on the Ewald sphere. Assuming that the phases of the diffraction pattern can be recovered from the measured intensities (for a discussion of the phase problem see Chapter 8), the 2D Fourier transformation of the diffraction pattern back to real space yields the correct two-dimensional object potential $\varphi(x, y)$. This result can also be obtained applying the Fraunhofer approximation to the classical Kirchhoff formula for a two-dimensional object. In this approach, which is described in detail for example in [55], the effect of the object on the amplitude and phase of the incident wave is described by multiplication with a two-dimensional, complex-valued transmission function $t(x, y)$.

In order to judge if a treatment of the object as a 2D-object is valid in our case, consider the following: an object of thickness D causes modulations of the Fourier amplitude in direction of q_z. The width of these speckles is $1/D$. If the Ewald sphere at the largest measured diffraction angle still cuts through the same speckle centred at $q_\mathrm{z} = 0$, the measurement is equivalent to a measurement on the plane $q_z = 0$ through the origin. Thus, if

$$q_\mathrm{z,\,max} \lesssim \frac{1}{2D}\,, \tag{7.19}$$

the object can be considered as two-dimensional and it can be retrieved by inverse two-dimensional Fourier transform of the diffraction pattern, again assuming that phase recovery is possible. For a wavelength of 1 Å it follows from (7.18) and (7.19), that an object that shall be imaged with atomic resolution, must be atomically thin to be considered as two-dimensional. Thus, a molecule of several nanometers in thickness, e.g. a DNA or a carbon nanotube, as it is used as a specimen in our case, must be treated as a three-dimensional object. Since the 2D-image reconstructed from a single 2D diffraction pattern is not a projection along the z-axis of this three-dimensional object either, it must be interpreted by means of the object exit wave, hence by means of classical scalar diffraction theory.

7.2.2 The object exit wave

The wavefield immediately after having passed the object, also denoted as *object exit wave*, can be described by multiplication of the incident wave function with kind of a transmission function $t(x, y)$. This function $t(x, y)$ contains the information about the changes in phase and amplitude of the initial wave function due to the whole three-dimensional object. Hence, it is in general a complex-valued function. According to the Fraunhofer approach mentioned above, the far-field diffraction pattern corresponds to the Fourier transform of the transmission function $t(x, y)$. Thus, assuming that the phases could be retrieved, $t(x, y)$ and with this the wavefield in the object exit plane can be obtained by inverse two-dimensional Fourier transformation of the diffraction pattern. The wavefield in other planes can then be calculated by propagating the object exit wave, i.e. by convolution with the appropriate Fresnel propagator [56]. For propagation by a distance Δz, the wave function must be convolved with

$$p(x, y) = \frac{\mathrm{i}}{R\lambda} \exp\left(\mathrm{i}\pi(x^2 + y^2)/\lambda\Delta z\right) \tag{7.20}$$

in real space [54], or equivalently, its Fourier transform must be multiplied by the corresponding propagator term in reciprocal space

$$P(q_\mathrm{x}, q_\mathrm{y}) = \exp\left(\mathrm{i}\pi\Delta z\lambda(q_\mathrm{x}^2 + q_\mathrm{y}^2)\right) . \tag{7.21}$$

The latter corresponds to assigning the phase shifts to the values in reciprocal space which correspond to a shift of position by Δz in real space. By propagating the wavefield through the planes of the object, the different object planes can be brought into focus, however there is always the out of focus signal of the other object planes superimposed. Nonetheless, in this way one can extract information about the object in z-direction, which is contained in the measured diffraction pattern due to the curvature of the Ewald sphere. The method of reconstructing a two-dimensional image from a diffraction pattern of a three-dimensional object by Fourier transformation and subsequent wavefield propagation has been demonstrated for simulated diffraction images [56] as well as for experimental x-ray diffraction patterns [44].

7.2.3 The two-dimensional object function

In order to make the approach based on the transmission function more comparable to the result obtained by the Born approximation, a two-dimensional object function $f(x, y)$ shall be introduced as

$$f(x, y) = t(x, y) - 1\ . \tag{7.22}$$

The object function is zero where the incident wave is not changed by the object. Thus for an object of limited size the support of the function $f(x, y)$ is, in contrast to the transmission function, also limited to a finite region. The Fourier transform of the transmission function can then be written as

$$\begin{aligned}\mathscr{F}\{t(x, y)\} &= \mathscr{F}\{1 + f(x, y)\} \\ &= \delta(q_{\mathrm{x}}, q_{\mathrm{y}}) + \mathscr{F}\{f(x, y)\}\ .\end{aligned} \tag{7.23}$$

Thus, the far-field diffraction pattern, except for the central spot due to the unscattered part of the wave contained in the δ-function, is given by the Fourier transform of the object function $f(x, y)$. In the case of a two-dimensional potential the formulation in (7.23) is equivalent to the result of the Born approximation, given by equations (7.10) and (7.13), with the two-dimensional potential function $\varphi(x, y)$ corresponding to the object function.

Chapter 8

The phase problem

When an object is illuminated with a plane wave, the far-field diffraction pattern corresponds to the Fourier transform of the object function containing the information about the scattering properties of the object as has been discussed in 7.2. However, only the intensity of the diffraction pattern can be recorded at the detector, whereas the phase is lost during the detection process. Thus, direct recovery of the object function by applying an inverse Fourier transform to the diffraction pattern is not possible. This fact is widely known as the phase problem. In the following, some general aspects of the phase problem will be discussed, thereby introducing the necessary mathematical tools. Subsequently, some differences of the phase problem as it is encountered in x-ray crystallography from the one in coherent diffraction microscopy will be pointed out and the conditions for a unique solution of the phase problem are discussed. A brief survey of existing algorithms for iterative phase retrieval from the measured diffraction intensities will be given. Since the aim of our experiment is the retrieval of the two-dimensional complex-valued object function from the measured diffraction intensities, the discussion will concentrate on the two-dimensional case. However, the collection of three-dimensional Fourier intensities by rotation of the object about an axis normal to the beam direction is in principle possible, and possible extensions of phase retrieval methods to three dimensions shall thus be kept in mind. In many cases the statements on the phase problem for the two-dimensional case given here can easily be extended to three dimensions. Otherwise, if differences between the two- and three-dimensional case exist, this will be mentioned explicitly.

8.1 General aspects and mathematical tools

The continuous Fourier transform of a complex-valued function $f(x, y)$, is defined as

$$\mathscr{F}\{f(x,y)\} = F(u,v) = \int_{-\infty}^{\infty} f(x,y)\, e^{-\mathrm{i}(q_\mathrm{x} x + q_\mathrm{y} y)} dx\, dy\ , \tag{8.1}$$

where $F(u, v)$ is also a complex-valued function and can be written as

$$F(u, v) = |F(u, v)| \cdot e^{\mathrm{i}\Phi(u,v)} . \tag{8.2}$$

The function $f(\mathbf{x})$ can be recovered by applying the corresponding inverse Fourier transform to $F(u, v)$:

$$\mathscr{F}^{-1}\{F(u, v)\} = f(x, y) = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} F(u, v)\, e^{\mathrm{i}(q_\mathrm{x} x + q_\mathrm{y} y)} dq_\mathrm{x}\, dq_\mathrm{y} \tag{8.3}$$

$$\tag{8.4}$$

During the detection process of an experimental diffraction pattern however, the phases $\Phi(u, v)$ of the diffraction pattern are lost, precluding direct recovery of the object function $f(x, y)$ by simply applying an inverse Fourier transform. If the object function can be an arbitrary complex-valued function without any constraints, any phase function $\Phi(u, v)$ can be chosen to be combined with $|F(u, v)|$, corresponding to infinitely many possible object functions. In practice, however, there is usually some a priori knowledge about the object, e.g the approximate size of the object is often known and thus the object function is known to be zero outside a finite region of support. Another very powerful constraint, valid for many x-ray applications, is that the function must be real and non-negative, since, at least away from absorption edges, it corresponds to the electron density within the specimen [54]. Besides, knowledge about the atomicity and symmetry properties of the object can often be used as constraints for phase retrieval methods in crystallography. Thus, the phase problem can also be understood as an inverse problem, which consists in finding an object function leading to the measured intensity values of the diffraction pattern and at the same time fulfilling additional constraints in the object domain. This inverse problem may have a solution which is uniquely determined except for some properties of the object function that are irretrievably lost due to the missing phase of the Fourier transform.

As the functions $f(\mathbf{x})$, $f^*(-\mathbf{x})$, $f(\mathbf{x}-\mathbf{x_0})$ and $e^{\mathrm{i}\phi} f(\mathbf{x})$ all have the same Fourier modulus ($*$ denoting the complex conjugate of a function), it is impossible to recover the position of the object and its absolute phase and one can not distinguish between the object and its complex conjugated and inverted version. However, these properties do not affect the appearance of the object. Thus, any of these versions can be regarded as physically equivalent and if one of them can be retrieved from the measured Fourier modulus, the above stated inverse problem can be regarded as solved uniquely in a physical sense. Nevertheless, these intrinsic ambiguities shall be kept in mind, as they can cause severe problems concerning the convergence of iterative reconstruction algorithms.

8.1.1 The autocorrelation function

As follows from the correlation theorem, the inverse Fourier transform of $\left|F(q_\mathrm{x}, q_\mathrm{y})\right|$ yields the autocorrelation function $\mathscr{A}(x, y)$ of the object function

$$\mathscr{F}^{-1} \left|F(q_\mathrm{x}, q_\mathrm{y})\right| = \mathscr{A}(x, y) , \tag{8.5}$$

where the autocorrelation is defined as

$$\mathscr{A}(x,y) = \int_{-\infty}^{\infty} f^*(\tilde{x},\tilde{y}) \cdot f(x+\tilde{x}, y+\tilde{y})\, d\tilde{x}\, d\tilde{y}\,. \tag{8.6}$$

If $f(\mathbf{x})$ is a function with finite region of support, e.g. it is zero outside a region $S \in \mathbb{R}^2$, then the region of support of its autocorrelation function has twice the dimensions in each spatial direction. For example, if the object function is zero outside a rectangular region

$$S = \{x, y \in \mathbb{R}; \quad |x| \le \frac{a}{2},\ |y| \le \frac{b}{2}; \quad a, b \in \mathbb{R}\} \tag{8.7}$$

then the support of the autocorrelation function is contained in the rectangle of twice the dimensions

$$S_1 = \{x, y \in \mathbb{R}; \quad |x| \le a,\ |y| \le b; \quad a, b \in \mathbb{R}\}\,. \tag{8.8}$$

Thus, information about the object size can directly be obtained from the measured intensity values of the diffraction pattern.

8.1.2 Discrete images and sampling

As the detection of intensity distributions of any kind of radiation as well as further data processing always goes along with discretization of the data, one important insight from the theory of signal processing shall be stated here:

> *Let $f(x)$ be a band limited function so that its Fourier Transform vanishes for $|u| \ge u_{max}$. The function $f(x)$ can be completely reconstructed from its values sampled at intervals of Δx provided that $\Delta x \le 1/2u_{max}$.*

This is known as the Nyquist-Shannon sampling theorem. As the relation between a function and its Fourier transform is symmetric, the theorem can also be formulated for the reverse case of the function itself having finite support, i.e. being ‘bandlimited’, and sampling its Fourier transform:

> *Let $f(x)$ be a function with finite support, e.g. $f(x) = 0$ for $|x| \ge a/2$. Its Fourier transform $F(u)$ can be completely reconstructed from its values sampled at intervals of Δu provided that $\Delta u \le 1/a$.*

Stated in this way, it can directly be applied to the case considered here i.e. the case of sampling the far-field diffraction pattern of a finite sized object. It can be extended to two and more dimensions just by defining a region of support as done in (8.7) and defining the two sampling rates in the corresponding directions respectively.

For simplicity, an object function with a quadratic region of support of size a will be assumed in the following. If it was possible to not only measure the intensity values of the diffraction pattern but also the phases, a sampling rate at the detector of $1/a$ corresponding to the reciprocal value of the support size would suffice for a reconstruction of the diffraction pattern as a continuous function and hence for retrieving the object function. When only the intensity values are measured, one obtains samples of the Fourier transform of the autocorrelation of the object function. According to the Nyquist-Shannon sampling theorem, for this continuous function to be reconstructable from its sampled values, the sampling rate must at least correspond to the reciprocal value of the extent of the autocorrelation function, thus to $1/2a$. Thus, if the diffraction pattern is sampled with a rate of $1/2a$ or higher, the modulus of the object function's Fourier transform and with this the autocorrelation function are available as continuous functions. Hence, mathematical theory for this class of functions can be used to tackle the phase problem.

8.2 The phase problem in crystallography

An ideal crystal consists of infinitely many copies of a unit cell, which are arranged periodically in three-dimensional space. If each unit cell is referenced to a point, all these points $\mathbf{x_n}$ form a regular three-dimensional lattice, which can be described by three lattice vectors $\mathbf{a_1}$, $\mathbf{a_2}$ and $\mathbf{a_3}$. The electron density in an ideal crystal can thus be described by the convolution of the electron density of a single unit cell $e(\mathbf{x})$ (being zero outside the unit cell) with an infinite train of three-dimensional delta functions, constituting the lattice function $g(\mathbf{x})$:

$$\begin{aligned} f(\mathbf{x}) &= e(\mathbf{x}) \otimes \sum_{n=-\infty}^{\infty} \delta^3(\mathbf{x} - \mathbf{x_n}) \\ &= e(\mathbf{x}) \otimes g(\mathbf{x}) \end{aligned} \tag{8.9}$$

An x-ray diffraction pattern $F(\mathbf{q})$ of an ideal crystal is given by the Fourier transform of the three-dimensional electron density as follows from the Born approximation given in 7.2. Applying the convolution theorem, this yields:

$$F(\mathbf{q}) = E(\mathbf{q}) \cdot G(\mathbf{q}) \,, \tag{8.10}$$

where the Fourier transforms of functions are denoted with capital letters respectively. For simplicity, the case of a cubic lattice shall be assumed here. The coordinates of a vector $\mathbf{x}$ with respect to the basis spanned by the orthogonal basis vectors of length a shall be referred to as (x, y, z). The lattice points can be written as $\mathbf{x_n} = (m, n, p)$ with $m, n, p \in \mathbb{Z}$. With this, the three-dimensional δ-function in (8.9) can be written as a product of one-dimensional δ-functions, yielding

$$g(\mathbf{x}) = g(x, y, z) = \sum_{m,n,p} \delta(x - m)\delta(y - n)\delta(z - p) \,. \tag{8.11}$$

Denoting the coordinates in reciprocal space by (u, v, w), the Fourier transform of the lattice function $g(\mathbf{q})$ is thus given by the reciprocal lattice

$$G(\mathbf{q}) = \frac{1}{a^3} \sum_{n=-\infty}^{\infty} \delta^3(\mathbf{q} - \mathbf{K_n}) \,, \tag{8.12}$$

with $\mathbf{K_n} = (h, k, l)$; $m, n, p \in \mathbb{Z}$ being the lattice points in reciprocal space. With this, the Fourier transform of the electron density can be calculated to

$$\begin{aligned} F(\mathbf{q}) &= F(u, v, w) \\ &= \frac{1}{V}\, E(u, v, w) \cdot \sum_{h,k,l} \delta\,(u - h)\,\delta\,(v - k)\,\delta\,(w - l) \\ &= \frac{1}{V}\, \sum_{n} E\,(\mathbf{q})\,\delta^3(\mathbf{q} - \mathbf{K_n})\,, \end{aligned} \tag{8.13}$$

where V is the volume of the unit cell. Thus, the diffraction pattern of an ideal crystal consists of discrete points, the Bragg peaks, which are samples of the Fourier transform $E((q))$ of the electron density of a single unit cell. The relation between the Fourier transform of a unit cell and the Fourier transform of this unit cell repeated regularly in two directions is also illustrated in Fig. 8.1.

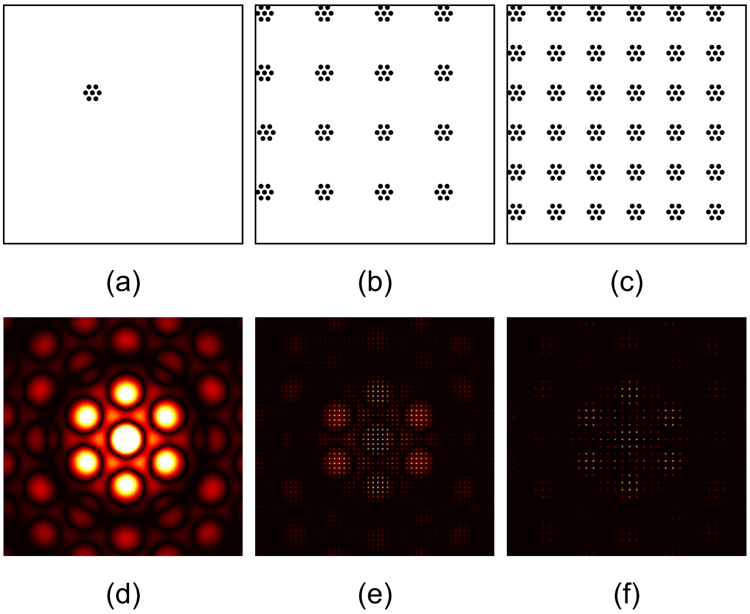


Figure 8.1: *Illustration of the relationship between periodic repetition of the object and the modulation of the diffraction pattern. (a) single non-periodic object and part of the amplitude of its continuous FT (d). If the same object is repeated periodically in both directions (b), the FT consists of discrete points (e), obviously being samples of the continuous function shown in (d). In (c) the distance between the individual objects is decreased, resulting in coarser sampling of the FT (f).*

Physically, the high intensity at the Bragg peaks results from constructive interference between the scattered waves originating from the copies of the unit cell in the crystal. For a finite crystal of N unit cells, the amplitude at these positions of constructive interference equals the N-fold of the amplitude of the diffraction pattern of one single unit cell only. Due to this gain of intensity, the Bragg peaks of a crystal are observable even though the brightness of the source would not suffice to obtain sufficient diffraction signal of a single unit cell. This fact made the experimental detection of Bragg peaks in x-ray crystallography possible at a time, when there were no high-brightness sources available. For an ideal crystal of infinite size, the intensity in between the Bragg peaks is zero. For a crystal of finite extent this is not exactly true, but for typical crystal sizes, the signal in between the Bragg spots is still too low to be detectable. Thus, the detected intensity distribution of a crystal's diffraction pattern is the sampled version of the intensity of the diffraction pattern of a single unit cell. The sampling rate is thereby intrinsically determined by the length of the reciprocal lattice vectors. In the case of a cubic lattice with lattice constant a, the sampling rate is $1/a$, or, similarly, for any other type of crystal system the sampling rate in each direction is the reciprocal value of the corresponding distance of the lattice planes. As stated above, the Nyquist frequency for adequate sampling of the diffraction intensity is the reciprocal of the size of the autocorrelation function. Here, it is thus given by the reciprocal of twice the size of the unit cell. Thus, in crystallography, the diffraction intensity of the unit cell is intrinsically undersampled by a factor of two. Hence, according to the Nyquist-Shannon theorem, the underlying continuous function $|E(\mathbf{q})|$, i.e. the modulus of the Fourier transform of the electron density inside a unit cell, can not be reconstructed from the detected samples.

8.3 The phase problem in coherent diffraction imaging

In contrary to x-ray crystallography, the diffraction pattern of a non-periodic object, as encountered in coherent diffraction imaging, is continuous. The sampling rate is not intrinsically limited but it is only determined by experimental conditions, namely by the detector properties and the spatial coherence of the radiation. In coherent diffraction imaging it is thus possible to sample the diffraction intensity on a sufficiently fine grid to allow the underlying continuous function to be retrieved according to the sampling theorem. For this reason, the phase problem in crystallography appears to be different from the one in coherent diffraction imaging. Concerning the uniqueness of the solution, approaches to the phase problem based on the availability of the modulus of the Fourier transform as a continuous function, do not apply to crystallography. Hence, the methods for phase retrieval in crystallography are quite different from the ones developed for coherent diffraction imaging, which will be discussed in the following sections. A thorough overview of differences and similarities of the phase problem in crystallography and in general imaging can be found for example in [57].

8.3.1 Uniqueness of the phase problem for non-periodic objects

The sampling rate of the diffraction pattern of a non-periodic object is not intrinsically limited. One can therefore assume that the sampling rate is higher than the Nyquist frequency and allows retrieval of the underlying continuous function. Thus, the question for the constraints that must be imposed on the object function for the phase problem to have a unique solution, can be tackled for the case of continuous functions. This has been done for two and more dimensions during the 1970 s and 1980 s [58, 59, 60, 61]. At the same time, first iterative algorithms for object retrieval became available. As it was already known before, there is in general no unique solution to the phase problem in the one-dimensional case [62]. However, numerical experiments, applying iterative phase retrieval algorithms to two-dimensional discrete images, suggested that the situation is by far more favourable for the two-dimensional case, since the algorithms converged to the correct object function [36]. This conjecture was confirmed theoretically for the case of continuous complex-valued two-dimensional functions by Barakat and Newsam in 1984 [59]. They could show, that the phase problem is almost always uniquely solvable in two dimensions if the object function has a support contained in a bounded subset of $\mathbb{R}^2$ except for the ambiguities mentioned above. Thus, for the realistic case of a finite sized object, the phases of the diffraction pattern are in general uniquely determined if the intensity distribution is known as a continuous function. For this, as stated above, the sampling rate must at least correspond to the reciprocal of twice the object size.

One can also choose a less mathematical approach to show that the complete information about the intensity of the diffraction pattern is only contained in its sampled version, if the sampling rate is larger then $1/2a$. For this, a one-dimensional object of size a shall be considered, as it is depicted in Fig. 8.2. The intensity of its far-field diffraction pattern shall be recorded at a detector placed at a distance R. The finest features in the diffraction pattern are due to interference of the waves scattered by the two outermost points of the object, since these correspond to the lowest frequency contained in the Fourier spectrum. The distance between the intensity maxima of the diffraction pattern of just these two points on the detector plane is $y = R\lambda/a$. For these intensity maxima to be distinguished in the recorded diffraction pattern, at least two samples per period are required according to the sampling theorem. Thus, the pixel size of the detector must be smaller than $R\lambda/2a$ which corresponds to a frequency in Fourier space of $\Delta u = 1/2a$, hence the same result as obtained above. This is the Nyquist-Shannon sampling rate for the intensity of the diffraction pattern and would correspond to two-fold oversampling of the complex-valued diffraction pattern.

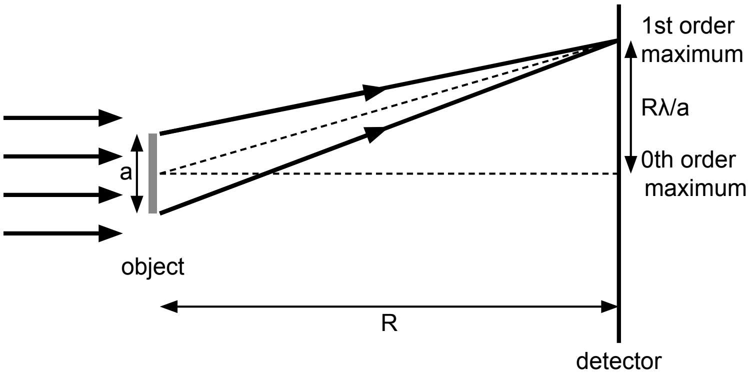


Figure 8.2: *Distance between the intensity maxima at a distant detector due to scattering of a plane wave on the two outermost points of an object of size a.*

In the literature about coherent diffraction microscopy, the expression *oversampling phasing method* has been established by Miao [63] for iterative phase retrieval from a diffraction pattern sampled with a rate higher than the reciprocal of the object size $1/a$. The term *oversampling* will therefore

be used accordingly in the following. An oversampling ratio O can be defined, giving the degree of oversampling with respect tot the Nyquist sampling rate $1/a$.
For the three-dimensional case, it has been proven [64], that the phase problem is overdetermined if the amplitude of the Fourier transform is known as a continuous function. This means that the three-dimensional Fourier amplitude oversampled by a factor of two in each direction contains redundant information. Therefore, sampling requirements can be somewhat relaxed for three-dimensional imaging, reducing the amount of data to be collected in an experiment.

8.4 Phase retrieval algorithms in coherent diffraction imaging

As discussed above, the phase information is encoded in the measured samples of the diffraction intensity for the case of finite-sized objects, provided the sampling rate meets the above stated requirement. It is then a matter of extracting this information from the measured data. To this end, several algorithms have been developed. Here, after a short introduction to the discrete Fourier transform, I will focus on the phase retrieval algorithms which have been invented in the 1970s by Fienup for applications in astronomical interferometry [36, 65] and since have been modified in order to improve their performance when applied to experimental coherent diffraction images. Some of these modifications will be sketched, concerning the reconstruction of complex-valued object functions and the problem of the missing low resolution data in coherent diffraction images. Additionally, some aspects of the influence of noise on the phase retrieval routines will briefly be discussed. The discussion will be restricted to the two-dimensional case, although an extension of the presented algorithms to three dimensional data sets is possible as has for example been shown by several authors [44, 48, 66].

8.4.1 Discrete Fourier transform and oversampling

When dealing with discrete images, one uses the discrete Fourier transform (DFT) instead of the continuous Fourier transform defined in (8.1). Assume a two-dimensional discrete image of size $M \times N$, with the pixel values being samples of a function of a continuous two-dimensional variable, taken at $x_m = m\,\Delta x$ and $y_n = m\,\Delta y$, $m = 0 \ldots M-1$ and $n = 0 \ldots N-1$. The DFT of the discrete function $f(x_m, y_n)$ is defined as

$$F\left(u_k, v_l\right) = \sum_{m=1}^{M} \sum_{n=1}^{N} f\left(x_m, y_n\right) e^{-\mathrm{i}\left(\frac{km}{M} + \frac{ln}{N}\right)} \,, \tag{8.14}$$

with $u_k = k\Delta u$, $v_l = l\Delta v$ and $k = -M/2 \ldots M/2, \quad l = -N/2 \ldots N/2$. The sampling intervals in Fourier space must satisfy the Nyquist sampling condition

$$\Delta u = \frac{1}{M\Delta x}\,, \qquad \Delta v = \frac{1}{N\Delta y}\,. \tag{8.15}$$

The function itself can be re-obtained by applying the corresponding inverse DFT

$$f\left(x_m, y_n\right) = \frac{1}{2\pi} \sum_{k=1}^{M} \sum_{l=1}^{N} F\left(u_k, v_l\right) e^{\mathrm{i}\left(\frac{km}{M} + \frac{ln}{N}\right)} \tag{8.16}$$

with $m = 0 \ldots N - 1$ and $n = 0 \ldots N - 1$.

The DFT is constructed such that the Nyquist sampling condition is exactly fulfilled in real space as well as in Fourier space, considering the size of the image as object size and $u_{\max} = \frac{N}{2}\Delta u$, respectively $v_{\max} = \frac{N}{2}\Delta v$ as band limits. If the sampling interval in Fourier space is reduced by increasing the number of pixels, the pixel sizes in real space Δx and Δy stay the same, according to (8.15). However, the size of the image obtained by an inverse Fourier transform is increased and the original image of size $M \times N$ is padded with zeros.

The phase problem for a continuous object with finite support, as discussed above, can now be looked at from a different point of view: when the intensity of a diffraction pattern is sampled with a frequency higher than $1/a$, where a denotes the object size, the size of the region obtained by an inverse Fourier transform is larger than the object size. In fact, for the phase problem to have a unique solution, the sampling frequency must be at least $1/\,2a$ as stated above. This corresponds to a region in real space of twice the object size in each direction. Thus, the available information about the object, i.e. that it has finite size, can easily be applied to the real space image as the pixels outside the region of support are known to be zero. This fact is also exploited by the iterative algorithms that will be described in the following section.

8.4.2 Iterative reconstruction algorithms

The phase retrieval algorithms described here are based on a modification of the Gerchberg-Saxton algorithm [35, 67], an algorithm for phase retrieval in the case that intensity images are measured in both, in the object as well as in the diffraction plane, as it can be done in electron microscopy. The Gerchberg-Saxton algorithm involves back and forth between real (object) space and Fourier (diffraction) space by Fourier transformation imposing the respective sets of constraints in each step, i.e. the measured intensities values in both domains. In this way, the phases can be recovered in an iterative manner.

Fienup's so called error-reduction (ER) algorithm [36, 65] is a generalization of the Gerchberg-Saxton algorithm in the sense, that the measured intensities can be replaced partially by other constraints. In the case of coherent diffraction imaging *a priori* information about the object is used as constraint in the object space, i.e. that the object function is zero outside a known region of support, since intensities are measured only in Fourier space. For some applications, as for the case of x-rays, the object function is additionally known to be real and non-negative.

The ER-algorithm starts by combining the measured values $|F(u, v)|$ with a guess of the phases in Fourier space $\Phi(u, v)$. For example a random phase distribution can be taken as a starting point. The subsequent iterative procedure consists of the following four steps which are also illustrated in Fig. 8.3 (a):

1. Inverse Fourier transformation back to the object space yielding a first estimate of the object

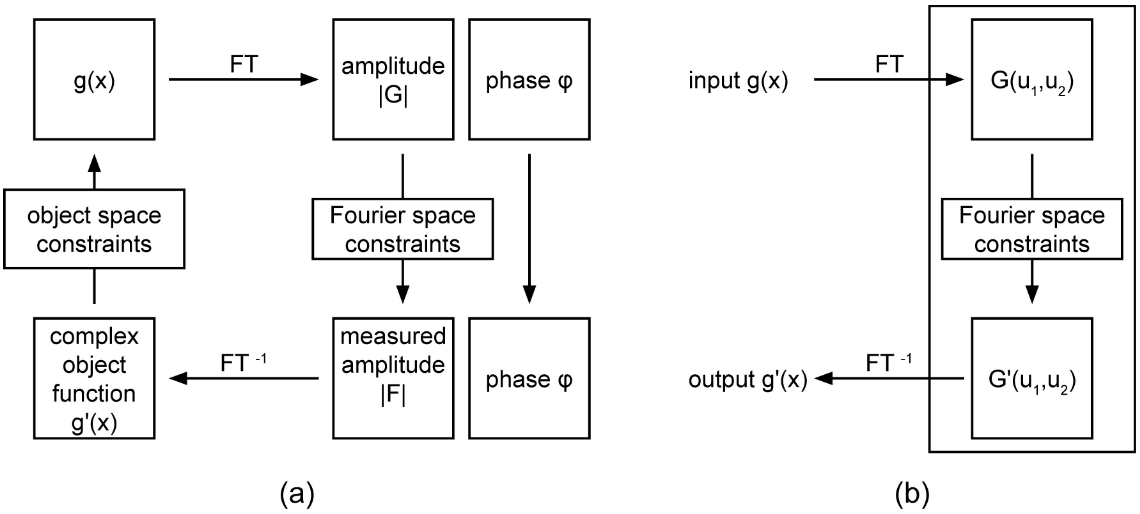


Figure 8.3: *Schemes of algorithms for phase retrieval. (a) Error-reduction algorithm (b) Hybrid input-output algorithm*

function

$$g'_{k+1}(x,y) = \mathscr{F}^{-1}\left\{G'_k(u,v)\right\} = \mathscr{F}^{-1}\left\{F(u,v)\cdot e^{\mathrm{i}\Phi_k(u,v)}\right\}\ . \tag{8.17}$$

2. The constraints in the object space are imposed on the estimate of the object function

$$g_{k+1}(x,y) = \begin{cases} g'_k(x,y) & \text{if } (x,y)\in\Omega \\ 0 & \text{if } (x,y)\notin\Omega \end{cases} \tag{8.18}$$

   where Ω denotes the region in $\mathbb{R}^2$ where the object constraints are fulfilled.

3. Fourier transformation of the object function yielding an updated estimate of phases in Fourier space

$$G_{k+1}(u,v) = \mathscr{F}\left\{g_{k+1}(x,y)\right\} = |G_{k+1}(u,v)|e^{\mathrm{i}\phi_{k+1}(u,v)}\ . \tag{8.19}$$

4. The amplitudes of the Fourier transform are replaced by the measured amplitudes

$$G'_{k+1}(u,v) = |F(u,v)|e^{\mathrm{i}\phi_{k+1}(u,v)}\ . \tag{8.20}$$

This scheme is recurred until the computed object function $g'_k(x,y)$ satisfies the object constraints and hence a Fourier transform pair has been found, fulfilling the constraints in object space as well as in Fourier space. When dealing with noisy images of course, it is sufficient that the constraints are fulfilled within a certain tolerance level. It can be shown, that the error-reduction algorithm converges, in the sense that the error in the object domain E_{Ok} and the error in the Fourier domain

E_{Fk}, which are defined as

$$E_{Ok}^2 = \frac{1}{N} \sum_{x \notin \Omega} |g_k'(x,y)|^2 \tag{8.21}$$

$$E_{Fk}^2 = \frac{1}{N} \sum_{u,v} |G_k(u,v) - G_k'(u,v)|^2 \tag{8.22}$$

$$\tag{8.23}$$

decrease or stay the same with every iteration step. The convergence however often becomes very slow after a few iterations [65]. To overcome this stagnation problem, the hybrid input-output algorithm (HIO) was developed [36, 65], which is a modification of the ER algorithm, offering some more flexibility in imposing the constraints in the object domain. For this, the diagram of the ER algorithm is no longer interpreted as a procedure to be passed through cyclically, but the steps at the right hand side of the diagram in Fig. 8.3 (a), or the steps (3), (4) and (1) in the above described procedure, are grouped together and considered as a non-linear system having an input $g_k(x,y)$ and an output $g_k'(x,y)$ (see also Fig. 8.3 (b)). The input of the next iteration $g_{k+1}(x,y)$ must no longer fulfil the object constraints, but it must be built such, that the next output $g_{k+1}'(x,y)$ is driven towards a solution. In a first approximation, the output depends linearly on the input, thus the output can be pushed in a desired direction by a small change of the input in the same direction. This procedure can be thought of as kind of a negative feedback loop. Thus, a logical choice of the next input appears to be

$$g_{k+1}(x,y) = \begin{cases} g_k'(x,y) & (x,y) \in \Omega \\ g_k(x,y) - \beta g_k'(x,y) & (x,y) \notin \Omega\,, \end{cases} \tag{8.24}$$

with Ω denoting the set of points at which $g_k'(x_1,x_2)$ satisfies the object constraints. The parameter β is typically between 0.5 and 1.0 and is often referred to as feedback parameter. It can be adjusted to optimize the convergence of the algorithm and can also be varied throughout the iterative procedure. It turns out, that the HIO algorithm in general converges much faster than the ER algorithm [36]. For reasons of stability a combination of both algorithms is often favourable in practice, performing a number of iterations of the HIO algorithm followed by several cycles of the ER algorithm [65].

8.4.3 Support constraints for the reconstruction of complex-valued objects

Initially, the described phase retrieval algorithms have been developed in order to reconstruct a real valued non-negative object function, as it can usually be assumed for x-ray diffraction as long as the object is thin enough that absorption can be neglected. In this case, apart from the constraint of finite support, the constraint of non-negativity can be applied in the object space. This reduces the degrees of freedom in the algorithm enormously and thus enhances the probability of the algorithm to converge to the correct solution even if the region of support is only approximately known and a loose support constraint is applied. In the more general case of a complex-valued object function, the loss of this powerful constraint of non-negativity must be compensated for the phase retrieval algorithm to converge nonetheless. If the only *a priori* information about the object is its size, this can only be done by choosing the support as tight as possible. The influence of the support constraint on the convergence of phase retrieval algorithms has first been studied by Fienup, especially for the reconstruction of complex-valued objects [68]. He pointed out that the reconstruction of complex-valued objects is *considerably more difficult*, however, the reconstruction turned out to be possible if the support constraint is tight enough or if the region of support has a special geometry. In practice, often only a loose support can be determined from the *a priori* knowledge about the size and the approximate shape of the object, and the phase retrieval algorithm may fail. Thus, instead of estimating the support once and then keeping it constant, it has been suggested to constantly adjust the region of support throughout the algorithm. This dynamical support is updated every few iterations basically by using a thresholded version of the current object estimate. A detailed description of such algorithms using a dynamical support can be found in [69, 70, 71], where a successful application to the reconstruction of complex-valued objects is also shown. A further improvement in the reconstruction of complex-valued object functions can be achieved by applying another modification of the HIO algorithm, which is known as charge-flipping algorithm [72, 73]. In many cases, best results are obtained by combinations of the different algorithms, alternating between them after a number of iterations.

8.4.4 Missing low resolution information

The information about the central low resolution part of the diffraction pattern around the zero-order beam is usually missing in experimental data due to overexposure of the CCD camera or due to covering of the non-diffracted beam by a beam-stop. The missing information is often compensated by filling in the central part with the Fourier transform of a low resolution image obtained by other available microscopy methods. This low resolution part of the Fourier transform can be kept fixed throughout the whole reconstruction procedure [37, 74] or, alternatively can be used only as a starting point, letting the values in the central region freely develop after a few iterations of the algorithm [45, 71].

Another approach consists in treating the pixels in the centre of the diffraction pattern as free parameters throughout the reconstruction procedure and let them be determined by the algorithm without adding any starting information [47, 69]. Several other methods exist, including modifications of the above mentioned ones, as they are described for example in [43, 71]. Of course, the best procedure for compensating the missing central part must be chosen according to the available data and the parameters must be adjusted from case to case.

Experimentally, the problem of the missing low resolution information in the data can partly be overcome by recording several images with different exposure times for every diffraction pattern and composing them to a high dynamic range image, which would then contain the information about the central part as well [43]. Alternatively, an absorption filter could be used instead of varying the exposure times, to record the central part of the diffraction pattern [41]. A first attempt to integrate an LCD display as a variable absorption filter into our experimental setup has been made [75]. In our case however, since the incoming beam is not perfectly parallel but slightly divergent, part of the low resolution information will always be overlaid with the non-diffracted beam.

8.4.5 Effects of noise in the diffraction images

Experimental diffraction patterns are always deteriorated by different kinds of noise. First, as the detection of intensity is a discrete process, Poisson noise is present in any measured diffraction pattern. Since the recording of the diffraction pattern usually involves a charged coupled device (CCD) detector, the dark current and the read out noise of the CCD chip must also be taken into account. Hence, it is important to know how the numerical phase retrieval methods discussed above perform when applied to noisy images.

It is obvious that the high resolution information gets lost, if the weak signal of the higher diffraction orders is buried in noise. In case the phase retrieval algorithm converges to a solution, the resulting image reconstructed from a noisy diffraction pattern is therefore blurred and fine features may not be recognizable any more. Besides, the inverse problem of phase retrieval must be understood in a different way if noise is present: As the measured intensity values are not the *true* values, but each value is distorted by an amount of noise, the aim is not to find an object function $f(x, y)$ such that the equation

$$|\mathscr{F}\{f(x,y)\}|^2 = I(u,v) \ , \tag{8.25}$$

with $I(\mathbf{q})$ denoting the measured intensities, is exactly fulfilled, which wouldn't even be possible in most of the cases. Instead, an estimate of the object function must be found, which is likely to yield the measured intensity distribution $I(u, v)$. This means, the Fourier transform of this estimate may deviate from the measured values within a confidence level depending on the amount of noise. Thus, the inverse problem is solved in a least-square sense, meaning that a function $\tilde{f}(x, y)$ must be found so that $\mathscr{F}\{\tilde{f}(\mathbf{x})\}$ is as close as possible to the measured intensity distribution:

$$\tilde{f}(x,y) = \underset{f(x,y)}{\operatorname{argmin}} \left| I(x,y) - |\mathscr{F}\{f(x,y)\}|^2 \right| \ . \tag{8.26}$$

However, depending on the noise level, this modified inverse problem may become ill-posed in the sense that there is no unique solution [76].

It has been suggested [74] that finer sampling of the diffraction pattern may help in this case to compensate the loss of information due to an increased level of noise. The robustness of iterative phase retrieval algorithms against noise has been studied by various groups for several modifications of the phase retrieval algorithm [36, 38, 63, 71, 74, 77] showing that an estimate of the object function can still be obtained even with a significant amount of noise present in the diffraction image, although, of course, with reduced resolution. However, a systematic theoretically sustained examination of the influence of noise on the algorithms' performance, also in connection with the sampling rate, is lacking.

Chapter 9

Limitations in coherent diffraction imaging

The resolution in coherent diffraction imaging is in principle only limited by the wavelength and the largest scattering angle recorded. However, this is only true if the incident beam is perfectly monochromatic and parallel which of course is never strictly fulfilled in a diffraction experiment. Hence, an experimental far-field diffraction pattern always deviates from the Fourier transform of the object function, obtained in the idealized case. The limitations thus imposed, can be described in terms of limited spatial and temporal coherence. Deviations from the idealized diffraction pattern due to limited spatial coherence may preclude subsequent phase retrieval, because the corresponding algorithms are based on the assumption that the diffraction pattern is related to the object function via Fourier transformation. Here, the requirements shall be deduced, that must be met to keep the deviations so small that they are not detected in the experiment. Finite temporal coherence on the other hand imposes limitations on the achievable resolution, and will subsequently be analysed.

9.1 Spatial coherence

In our experimental setup the beam impinging on the sample is slightly divergent due to lens aberrations. It can be described as an incoherent superposition of plane waves with different incidence angles θ. The far-field diffraction patterns resulting from each of these plane waves are identical but slightly shifted relative to each other. Their incoherent superposition thus causes the features in the observed diffraction pattern to appear smeared out.

In order to quantify the effect of the beam divergence angle, a Young's double slit with a slit spacing a shall be considered as an object, illuminated by a monochromatic beam with half divergence angle θ as illustrated in Fig. 9.1. A detector is placed at a distance R from the object. The diffraction pattern at the detector due to scattering of the plane wave travelling parallel to the optical axis has its intensity maxima at the diffraction angles $\alpha = m\lambda/a$ with $m = 0, \pm 1, \pm 2, \dots$ and its minima at $\alpha = n\lambda/a$ with $n = \pm\frac{1}{2}, \pm\frac{3}{2}, \pm\frac{5}{2}, \dots$, where the small angle approximation has been applied. If the object is illuminated by a plane wave under an angle θ there is already an initial phase shift between the two waves emanating from the two slits. This leads to a shifted diffraction pattern with maxima at the angles $\alpha = m\lambda/a + \theta$. For the interference fringes still to be visible if these two diffraction patterns are superimposed, their relative shift must be smaller than half the fringe spacing. If the maxima of the diffraction pattern resulting from the tilted beam coincide with the minima of the pattern due to the beam parallel to the optical axis, the fringes are completely smeared out. This is the case if

$$\theta = \frac{\lambda}{2a} \,. \tag{9.1}$$

Accordingly, for a given beam divergence angle one defines the transverse coherence length ξ_t in the sample plane as

$$\xi_t = \frac{\lambda}{2\theta} \,. \tag{9.2}$$

Hence, as long as the object size is smaller than the coherence length, all the features in the diffraction pattern are still visible. It is, however, obvious that the diffraction pattern nonetheless exhibits deviations from the idealized one that would be obtained with a perfectly parallel beam. Approaches for phase retrieval in coherent diffraction imaging, relying on the assumption that the object function and the diffraction pattern are connected via Fourier transformation, are only valid if these deviations are not *noted* by the detector. Therefore, the requirement on the beam divergence must be related to the detector properties, in particular to the pixel size of the detector. Hence, it becomes somewhat more rigorous than stated above in (9.1), as shall be deduced in the following:

If the diffraction pattern is recorded in far-field at a detector with a pixel size Δy, every pixel covers an angle $\Delta\alpha = \Delta y/R$ corresponding to a frequency interval $\Delta u = \Delta\alpha/\lambda$. If the diffraction pattern of an object of size a shall be oversampled by a factor O, the pixel size must correspond to an angle $\Delta\alpha = \lambda/Oa$. Considering again the Young's double slit experiment described above, the angular shift between the interference patterns caused by the two beams with maximal divergence angle of $\pm\theta$ must be less than half the angle spanned by one pixel in order not to be detected. Thus, the condition

$$\theta \leq \frac{\lambda}{2\,Oa} \tag{9.3}$$

must be fulfilled. This means, according to (9.2), the transverse coherence length ξ_t must be larger than the object size multiplied with the oversampling factor, or, one can also say, the entire region Oa which will be reconstructed from the oversampled diffraction pattern by a phase retrieval algorithm must be illuminated coherently and hence:

$$\xi_t \geq Oa \,. \tag{9.4}$$

This result may seem to be somewhat counter-intuitive at first glance, since the region outside the object support, whether it is a zero-density region or a *known* density region, i.e. an opaque mask, does not contribute to the diffraction pattern except for the zero-order spot. The result in (9.4) however is a mere consequence of considerations of the spatial resolution of the detector.

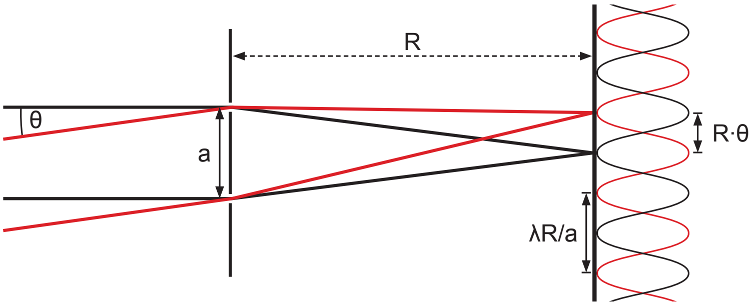


Figure 9.1: *Spatial coherence. If a Young's double slit is illuminated by a monochromatic beam with a divergence angle θ two diffraction patterns shifted to one another by an amount of $R\theta$ are superimposed at the detector at a distance R.*

9.2 Temporal coherence

Temporal coherence is a measure for the degree of coherence of the radiation along its propagation direction. It is therefore also often referred to as longitudinal coherence.

Consider two planar wavefronts with slightly different wavelengths, λ and $\lambda + \Delta\lambda$. If they are in phase at some starting point, they run out of phase while propagating freely in space and after a certain distance, defined as the longitudinal coherence length $\xi_l \approx N\lambda$ they have a phase difference of π. Then it is

$$N\lambda = (N - \frac{1}{2})(\lambda + \Delta\lambda) \tag{9.5}$$

and it follows

$$\xi_l \approx \frac{1}{2}\frac{\lambda^2}{\Delta\lambda}\,. \tag{9.6}$$

For an electron beam this can also be expressed in terms of energies as

$$\xi_l \approx \lambda\frac{E}{\Delta E}\,. \tag{9.7}$$

If in a diffraction experiment with a given energy uncertainty ΔE the path difference between the waves scattered off two different points of the object becomes larger than ξ_l, coherence is lost and the interference fringes disappear. In order to deduce the consequences on the resolution in the diffraction pattern, consider two planar wavefronts of slightly different wavelengths being diffracted by an object of size a under an angle α. The maximal path length difference for a certain diffraction angle occurs for waves scattered off the two outermost points of the object and is given by $\Delta = a\,\sin\alpha$. Assume that the interference pattern caused by the scattering of the wavefront with wavelength λ at the two outermost points of the object has its maximum of order N under the angle α. If the coherence length equals the path length difference, i.e. $\xi_l = \Delta$, this intensity maximum coincides with the N^{th} minimum of the interference pattern produced by the wave with wavelength λ+$\Delta\lambda$ and the visibility of the fringes in the superimposed interference pattern is reduced. Thus, for a given object size, the maximal angle where the beams of different wavelengths do add coherently is limited by the longitudinal coherence length ξ_l. Since the maximal diffraction angle α is related to the resolution d in the diffraction pattern by $d = \lambda / \tan\alpha$, the maximal achievable resolution is limited by the longitudinal coherence length given by

$$\xi_l \geq \frac{a\lambda}{d} \tag{9.8}$$

using $\sin\alpha \approx \tan\alpha$. This can also be expressed as a condition for the energy spread, yielding for the case of an electron beam:

$$\frac{E}{\Delta E} \geq \frac{a}{d} \tag{9.9}$$

For an experiment of coherent diffraction where the sampling of the diffraction pattern is considerably finer than half the fringe spacing, the changes induced to the diffraction pattern by the finite longitudinal coherence length, are measurable even though the condition in 9.8 is still fulfilled. It is not sufficient that the interference maxima due to the wave with wavelength λ and the one with wavelength $\lambda + \Delta\lambda$ are shifted with respect to each other by less than half the fringe spacing. Instead, the shift is required to be less than half a pixel. This leads to the condition for the energy spread in an oversampling experiment with the oversampling ratio O [78]:

$$\frac{E}{\Delta E} \geq \frac{Oa}{2d} \, . \tag{9.10}$$

Chapter 10

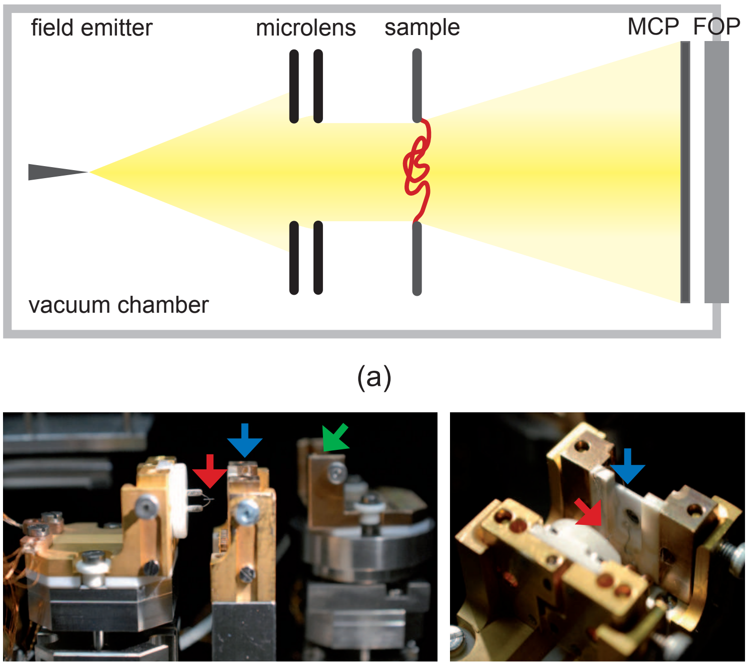
The experimental setup

10.1 The overall experimental scheme

A dedicated ultra-high vacuum system has been set up for coherent diffraction of low-energy electrons at single molecules. A divergent beam of coherent low-energy electrons emitted by a tungsten field-emitter is first collimated by an electrostatic microlens and then diffracted by the sample. The diffraction pattern is recorded at a detector placed at a distance of 75 mm from the sample. For detection, the electrons are multiplied by a microchannel plate (MCP) and then accelerated towards a phosphorescent screen. The latter is held at a voltage of several kilovolts so that the kinetic energy of the electrons is high enough for excitation of the phosphorescent coating and subsequent light emission. The phosphorescent coating is deposited on a fibre optic plate (FOP), guiding the emitted light from the vacuum to the ambient pressure side where the image is finally recorded by a CCD camera.

The field-emitter tip is placed at a distance of a few micrometers in front of the first aperture of the microlens, resulting in kinetic energies of the electrons at the entrance side of the lens of typically around 100 eV. For the alignment of the microlens and the electron source, the latter is mounted onto an x-y-z piezo positioning device which can be moved with nanometer precision in all three directions within an overall travel of 5 mm. The sample is mounted onto another x-y-z piezo-stage, being also rotatable around the vertical axis. This offers the possibility to illuminate the sample from different angles and thus to collect information about its three-dimensional structure. The whole sample stage can be moved aside, out of the beam, if this is needed for other than diffraction experiments.

Experiments are carried out at room temperature under ultra-high vacuum conditions ($5 \cdot 10^{-10}$ mbar). In order to reduce vibrations, the system is mounted on a separate platform, mechanically decoupled from the building and additionally air damped. A load-lock allows rapid transfer of microlenses and samples into the vacuum chamber. The system is equipped with a leak-valve and a diffusion valve for controlled introduction of neon or other noble gases, needed for sample cleaning and tip sharpening.



(a)

(b) (c)

Figure 10.1: *(a) Scheme of the experimental setup. The divergent coherent electron beam emitted by the electron source is first collimated into a parallel beam by the electrostatic microlens and then diffracted by the sample. For the detection of the diffraction pattern, the electrons are multiplied by a micro channel plate (MCP) and then reaccelerated towards a phosphorous screen on a fibre optic plate (FOP). The latter transfers the emitted light from the vacuum to the ambient pressure side, where it is captured by a CCD camera. (b) and (c) Two views into the vacuum chamber showing the electron point source (red arrow) mounted on a x-y-z piezo-stage, the microlens (blue) mounted on a dedicated holder, and the movable sample stage (green arrow), which can be pulled sideways into the optical axis.*

10.2 Electron source

Electrons are extracted by field emission from a sharp W(111) tip held at negative potential. A divergent beam of electrons of high spatial and temporal coherence is thus provided. In the following, a brief description of the principle of field emission shall be given, also including an estimate for the width of the energy distribution of the emitted electrons, determining the temporal coherence.

This is followed by a discussion of the electron optics of the initial cathode region, in order to show that in many cases, the field emitter can be treated as a point source. The limitations of this approximation will also be pointed out. The preparation method for the emitter tips employed in our setup is described at the end of this section.

10.2.1 Principle of field emission

Field emission is understood as the emission of electrons from the surface of a condensed phase into vacuum induced by a strong external electrostatic field. Electron field emission from a metal typically occurs at field strengths in the range of 10^9 V/m [28]. In order to achieve such high fields using reasonable potential differences, the emitter is shaped into a tip with an apex radius of curvature ranging from several micrometers down to a few nanometers. The potential barrier at the metal-vacuum interface is deformed by the external field as shown in the energy-diagram in Fig. 10.2. If the external field is high enough, the barrier becomes lower and sufficiently narrow for the electrons to have a significant probability for tunnelling from the Fermi edge of the metal into the vacuum. Thus, field emission is a purely quantum mechanical process. This is in contrast to thermionic emission or photoemission of electrons from a metal, where only electrons having sufficient kinetic energy may overcome the potential barrier.

For a theoretical treatment, the problem can be considered one-dimensional, which is a valid approximation, if the width of the potential barrier is small compared to the tip radius as it is the case for conventional field emitter tips and an external field of typically 10^9 V/m. The tunnelling probability for a particle with kinetic energy E and mass m through a one-dimensional potential barrier $V(x)$ of length L can be calculated using the Wentzel-Kramer-Brillouin approximation. It is given by [79]:

$$P(E,V) = f(E,V)\exp\left[\,-\,2\left(2m/\hbar^2\right)^{\frac{1}{2}}\int_0^L (V-E)^{\frac{1}{2}}dx\right], \tag{10.1}$$

where $f(E,V)$ is a function weakly depending on E and V [28].

The potential energy of the electron $V(x)$ outside the metal is a superposition of the work function barrier, which can be described as a step function, modified by an image force potential $-e^2/4x$, and the external electrostatic potential, as it is also depicted in Fig. 10.2. Taking the zero point of the potential at the vacuum level, this yields for an applied electric field F:

$$V(x) = -Fex - \frac{e^2}{4x} \tag{10.2}$$

The total tunnel current density is then given by the equation

$$j = e\int_0^\infty n(E_x)P(E_x,F)dE_x\,, \tag{10.3}$$

where $n(E_x)$ is the number of electrons with energies between E_x and $E_x + dE_x$ incident per unit time on a surface element of the barrier from within the metal; E_x is the part of the electron kinetic energy carried by the momentum component normal to the barrier surface. For the case of zero temperature the integration limits in (10.3) become $0 \le E_x \le \mu$ and the result is known as the Fowler-Nordheim equation [80, 81, 82]. With a correction factor $\alpha(e^3F)^{\frac{1}{2}}/\Phi)$ accounting for the

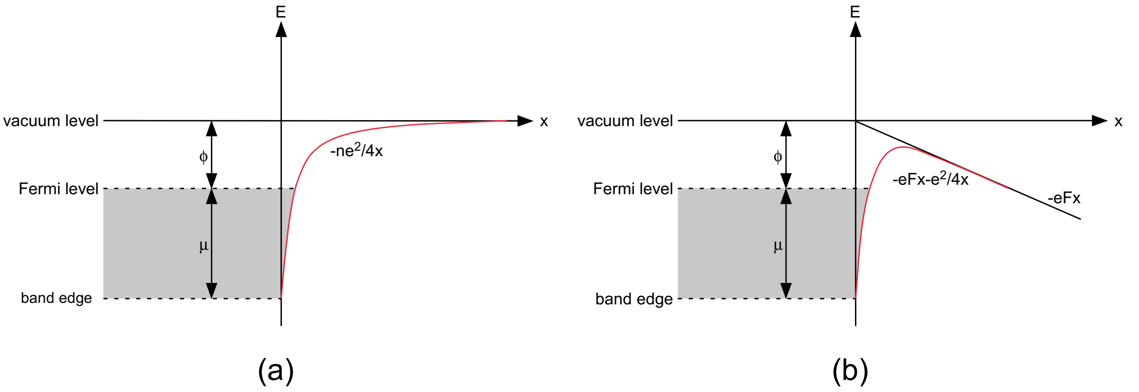


Figure 10.2: *Diagram of the potential energy of an electron at a metal-vacuum interface as a function of the distance x without (a) and with (b) external electric field.* $-e^2/4x$ *is the image force potential,* $-eFx$ *is the external applied potential,* Φ *is work function and* μ *the Fermi energy.*

image potential term in (10.2) the total current density is [28]:

$$j = \frac{e^3}{2\pi h}\frac{(\mu/\Phi)^{\frac{1}{2}}}{(\Phi+\mu)\alpha^2}F^2 \exp\left(-\alpha\frac{4}{3}\sqrt{\frac{2m}{\hbar^2}}\frac{\Phi^{\frac{3}{2}}}{eF}\right) \tag{10.4}$$

The image charge correction factor becomes $\alpha = 1$ if a sharp potential step is assumed at the metal surface instead of the hyperbolic increase due to the image charge potential. Tabulated values of the correction factor α are provided for example in [83].

The energy distribution of the emitted electrons $J(E)dE$ is important with respect to the temporal coherence of the beam and with respect to chromatic aberrations induced by any electron optical components. An expression for the distribution of the total energy of the electrons after tunnelling through the potential barrier has been derived by Young [84]. It is obtained by multiplying the incidence rate of electrons with total energy E on a surface element per unit time $J(E, E_x)dEdE_x$ with the tunnelling probability $P(E, V)$ and subsequent integration over the kinetic energy normal to the surface dE_x. For moderate temperatures one obtains [28]:

$$\begin{aligned} J(E)dE &= \int_0^E J(E, E_x)P(x)\, dE\, dE_x \\ &= \frac{32\pi m\mu^{\frac{1}{2}}F \exp\left(\frac{1}{2}b\,\Phi^{\frac{3}{2}}/F\right)}{3h^3 b\,(\Phi+\mu)} \cdot \frac{\exp\left(\frac{3}{2}b\,\Phi^{\frac{1}{2}}E/F\right)}{1+\exp\left[(E+\Phi)/kT\right]}\, dE\,, \end{aligned} \tag{10.5}$$

where b contains the correction term for the image effect:

$$b = \frac{-4\sqrt{2me}}{3\hbar}\alpha\,. \tag{10.6}$$

Fig. 10.3 (a) shows calculated values of the energy distribution at various temperatures. Beside the distribution of the total kinetic energy as calculated from (10.5), the so called normal-energy distribution for $T = 0K$ is also shown. This is the distribution of the energy associated with the momentum component normal to the surface immediately after tunnelling. Except for special electrode geometries, it becomes meaningless for electrons far from the emitting surface. The total-energy distribution in contrast also represents the distribution in total kinetic energy of all the electrons brought to a single potential anywhere outside the metal. The steep descent of the total-energy distribution at higher energies is due to the beginning of the Fermi sea and given by the Fermi-Dirac function contained in (10.5). It is smeared out at higher temperatures. At lower energies, the energy distribution drops exponentially according to the decrease of the tunnelling probability. The half width σ of the total-energy distribution can be calculated for the case of $T = 0K$ to [28]

$$\sigma = 0.46\frac{F}{b} \tag{10.7}$$

with F in V/m and b defined in (10.6); it increases slightly for higher temperatures.

Experimentally, the energy distribution $J(E)$ could be measured by Young and Müller for various temperatures [85]. A comparison between theoretical and experimental energy distributions for electrons emitted at room temperature is shown in Fig. 10.3 (b) confirming the theoretical results. For a field strength of $F = 5 \cdot 10^9 V/m$ at the field emitter tip (10.7) yields a width of the energy distribution of around 0.3 eV, which is also in accordance with the measured values displayed in Fig. 10.3 (b). If the electrons are extracted from the tip with a negative voltage of about 100 V, this results in a energy spread of $\Delta E/E \approx 3 \cdot 10^{-3}$.

The energy spread of the field emitted electrons on the one hand affects the chromatic aberrations when the electron beam is shaped by an electron lens. For the case of an electrostatic microlens as it is implemented in our experiment and as it has been described in Part I, chromatic aberrations due to an energy spread in the range of 10^{-3} turned out to be negligible compared to spherical aberrations (compare 4.3.1). On the other hand, an energy spread as small as possible is essential for coherent diffraction imaging in terms of temporal coherence, as it has been pointed out in 9.2. The energy spread deduced here, allows objects with a size of a few hundred nanometers to be imaged with a resolution in the Ångstrom range.

10.2.2 Virtual source size

As follows from the description of the field-emission process in the previous section, electrons are ejected from the tip apex, where the radius of curvature is small enough to yield a sufficiently strong electric field at the metal surface if a voltage is applied. Thus, electrons are emitted from an area corresponding to the apex radius R_0, which is in the order of 10 nm for the tips prepared as described in 10.2.3. In the region between tip and anode, the emitted electrons are accelerated by the electric field, such that their trajectories at the anode seem to emanate from a small region behind the physical tip apex, the *virtual source*, as it is schematically shown in Fig. 10.4. Thus, in many cases it is a sufficient approximation to treat the field emitter as a point source, emitting electrons with an energy corresponding to the accelerating voltage between tip and extractor electrode. This is often done when the field emitter is part of an electron-optical system. In order to determine the limitations of this approach, an estimate of the size of the virtual source will be given here, considering the electron-optics of the region between the tip and a planar extractor

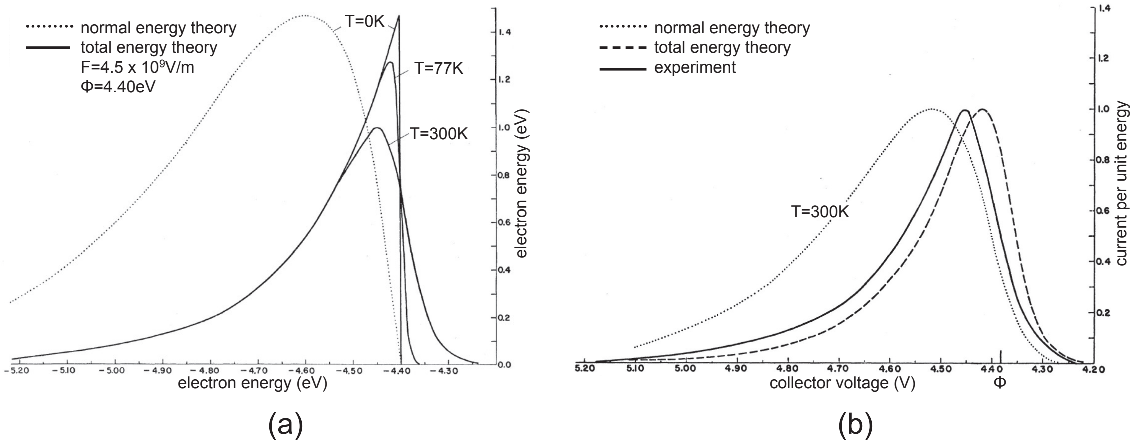


Figure 10.3: *(a) Theoretical total- and normal-energy distributions for field emitted electrons. The* 300 *K curve is arbitrarily normalized. (taken from Young [84]) (b) Comparison of theoretical and experimental total-energy distributions at room temperature for* $F = 3 \cdot 10^9\ V/m$ *and* $\Phi = 4.40\,eV$*. The theoretical normal-energy distribution is also shown. All curves are arbitrarily normalized. The collector voltage corresponding to the electron energy* $E = -\Phi$ *is labelled. (taken from Young et al. [85])*

electrode. The exact size of the virtual source is also crucial with respect to the spatial coherence of the emitted divergent electron beam. This can be exploited for experimental determination of the virtual source size.

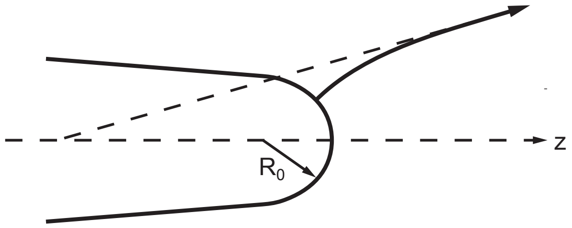


Figure 10.4: *Electrons emitted from the tip are deflected in the accelerating field between the tip and the anode such that their trajectories at the anode seem to emanate from a virtual source behind the tip apex. The geometry of the tip can be approximated by a sphere with a radius* R_0 *on a rotationally symmetric shank.*

Electron optics of the cathode region

Having just passed the tunnelling barrier, the electrons have Fermi energy. This initial kinetic energy of the emitted electrons is associated with velocity components in direction perpendicular as well as tangential to the emitter surface. The electron trajectories in the accelerating field between the tip and anode depend of course strongly on the geometry of the electrodes involved, i.e. on the exact shape of the apex as well as of the shank of the tip. The extractor electrode is assumed to be planar in the following. A simple model often used for the cathode region is the approximation of the tip by a sphere, fitting to the radius of curvature R_0 of the real tip apex (Fig. 10.4) and sitting on a rotationally symmetric shank. This model provides a good description of the field next to the apex showing only very small deviations from spherical symmetry. At larger distances from the tip, the real electric field is not at all spherical. However, in this region, since most of the potential drops in the vicinity of the apex, the electrons have already gained so much kinetic energy that the curvatures of the trajectories are very small and slight changes of the exact electric field direction have only little influence. The action of the electric field in this region can thus be regarded equivalent to a weak round lens, which can be described by basic tools of electron optical theory. The electron trajectories are thus first subject to a strong spherical field and then, at larger distances from the apex, they are further shaped according to the action of a weak round lens.

Spherical field in the vicinity of the tip apex For the effect of the spherical field next to the apex on the trajectories, consider an electron starting from the spherical surface with radius R_0 on the z-axis with an initial kinetic energy eV_0 and an initial elevation angle to the z-axis α as depicted in Fig. 10.5. Its trajectory can be calculated by applying the laws of conservation of energy and angular momentum. Using cylindrical coordinates, the electron's position at an equipotential corresponding to a voltage V_1 at distance R_1 from the centre of the assumed sphere is described by an elevation angle φ; the slope of its trajectory can be written as $\beta = \varphi + \Psi$, where Ψ is the angle from the trajectory's tangent to the radius vector in the point (R_1, φ) (see Fig. 10.5). At some distance $R_1 \gg R_0$ from the emitting surface, the electron trajectory will be almost perpendicular to the spherical equipotentials, since it is then $V_1 \gg V_0$. Thus, the following approximations for the angles are valid:

$$\varphi \ll \pi/2\,; \quad \Psi \ll \pi/2\,; \quad \beta \ll \pi/2\,, \tag{10.8}$$

and one can set

$$\sin\varphi = \varphi; \quad \cos\varphi = 1 - \frac{\varphi^2}{2}\,. \tag{10.9}$$

With these assumptions one obtains [86]

$$\varphi = 2\left(1 - \frac{R_0}{R_1}\right)\sin\alpha\sqrt{\frac{V_0}{V_1}} \tag{10.10}$$

$$\Psi = \frac{R_0}{R_1}\sin\alpha\sqrt{\frac{V_0}{V_1}}\,, \tag{10.11}$$

and with this for the slope of the trajectory

$$\beta = \varphi + \Psi = \left(2 - \frac{R_0}{R_1}\right)\sqrt{\frac{V_0}{V_1}}\sin\alpha\,. \tag{10.12}$$

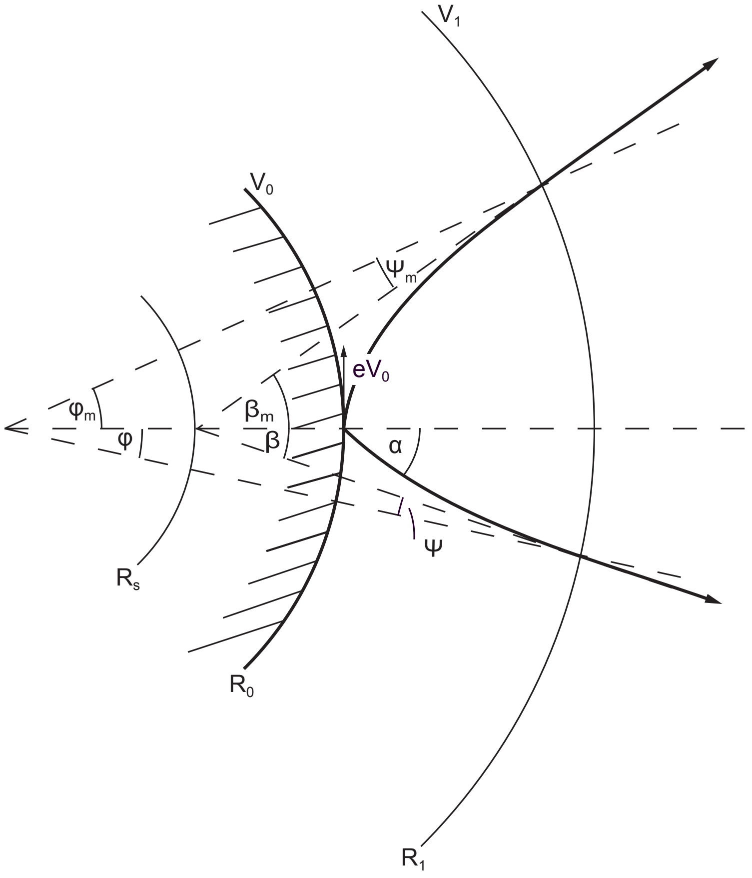


Figure 10.5: *Schematic drawing of the trajectories of two electrons starting with different initial elevation angles from a spherical surface with radius R_0. At a larger distance from the surface, the trajectories seem to emanate from a point on a sphere with radius R_S.*

The point of intersection of the trajectory's tangent in the point (R_1, φ) with the z-axis is given by

$$R_S = \frac{R_0}{2 - R_0/R_1} \,. \tag{10.13}$$

Since this does neither depend on the initial angle α nor on the voltage ratio V_1/V_0 for the applied approximations, all electrons starting from one point at (R_0, φ_0) with arbitrary starting angle and arbitrary initial kinetic energy, seem to emanate from a point (R_S, φ_0) behind the emitting surface. Thus, the emitting surface is imaged to a demagnified virtual surface having the radius R_S, as shown in Fig. 10.5.

The size of the virtual source however is given by the smallest cross section of the ray bundle extended backwards as shown in Fig. 10.6. For geometrical reasons, this virtual source is located approximately at the centre of the spherical surfaces. Its size can be determined by back propagation of the trajectories with maximal starting angle $\alpha = \pi/2$ to the centre of the spheres at $z = 0$

[87]. The elevation angle of theses trajectories at R_1 is given by (10.12) to

$$\beta_m = \left(2 - \frac{R_0}{R_1}\right)\sqrt{\frac{V_0}{V_1}} \approx 2\sqrt{\frac{V_0}{V_1}} \tag{10.14}$$

and thus the virtual source size is

$$r_0 = R_S \cdot \beta_m = R_0\sqrt{\frac{V_0}{V_1}}\,, \tag{10.15}$$

where the approximation $R_S \approx R_0/2$ has been used, valid for $R_1 \gg R_0$ according to (10.13). The voltage V_0 corresponds to the kinetic energy of the electrons immediately after the tunnelling process. Its mean squared value can be calculated from the total energy distribution and is given by [27]:

$$\langle V_0\rangle = \frac{e\hbar F}{2\sqrt{2m\Phi}} = \frac{1}{2}eF\lambda\!\!\!^{-} \quad \text{with} \quad \lambda\!\!\!^{-} = \frac{\hbar}{\sqrt{2m\Phi}}\,, \tag{10.16}$$

where the correction factor for the image force has been set to 1. With this, the mean square radius of the virtual source $\rho_0 = \langle r_0^2\rangle^{1/2}$ becomes

$$\rho_0 = R_0\sqrt{\frac{\langle V_0\rangle}{V_1}} = \sqrt{\frac{R_0\lambda\!\!\!^{-}}{2}}\,. \tag{10.17}$$

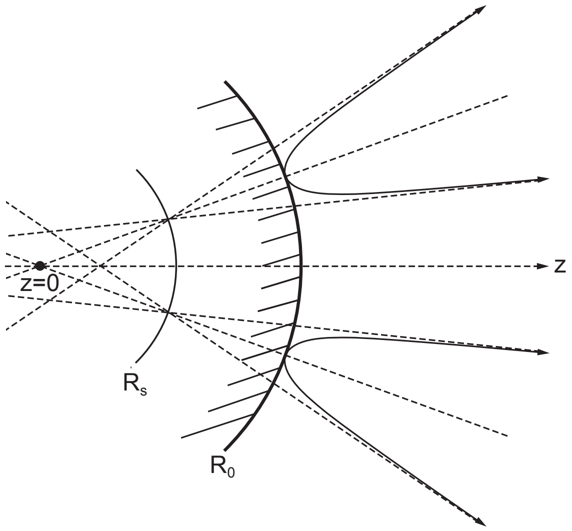


Figure 10.6: Bundle of trajectories extended backwards. The size of the virtual source is given by the smallest cross section through the back propagated ray bundle. It is found approximately at z = 0.

Field of a weak round lens At larger distances from the tip apex, the electron trajectories undergo the action of a weak round lens, further reducing the divergence angle of the trajectories.

The electrons entering this weak lens field seem to emanate from the virtual source as deduced before, i.e. from a region of size ρ_0. The elevation angle β of their trajectory is reduced by the accelerating lens field to an angle β_e at the anode according to

$$\frac{\sin\beta}{\sin\beta_e} = M_A\,, \tag{10.18}$$

where the value for the angular magnification has been calculated by numerical trajectory calculations for typical geometries to $M_A \approx 3^{-1/2}$ [27]. The refractive index defined for electrostatic electron optical systems as

$$n = \sqrt{\frac{V_1}{V_0}}\,, \tag{10.19}$$

denoting with V_i the voltages in the respective regions, in our case becomes

$$n = \sqrt{\frac{V_e}{R_1 F_1}} =: \sqrt{\mu}\,. \tag{10.20}$$

The fraction $V_e/(R_1 F_1)$ gives the ratio of the electron energy at the extractor electrode to the energy the electrons have already gained after passing the purely spherical field in the vicinity to the tip. Thus, the quantity μ turns out to be a purely geometric shape factor of the device [27]. With this, the lateral magnification of the weak lens field is given by [32].

$$M_L = \frac{1}{M_A\sqrt{\mu}} \approx \sqrt{\frac{3}{\mu}}\,. \tag{10.21}$$

Thus, the size of the virtual source ρ_0 as given in (10.17) is additionally changed by the action of the accelerating field further away from the tip apex to

$$\rho_v = M_L \rho_0 = \sqrt{\frac{3}{\mu}}\rho_0\,. \tag{10.22}$$

With $2.5 \leq \mu \leq 5$ [27], this does not differ much from ρ_0. It is essentially determined by the work function Φ and decreases with $1/\sqrt{R_0}$. The virtual source size is thus expected to be significantly smaller than the physical radius of curvature of the tip. For a tungsten emitter with an apex radius between 10 and 20 nm (10.17) yields values for the virtual source size of several Ångstroms.

Experimental determination of the virtual source size

In order to actually measure the virtual source size for field emitter tips similar to the ones used in our setup, several experiments have been carried out in the past. A divergent beam of electrons with an energy between 60 and 300 eV emitted from the tip is used to produce an interference pattern at a distant detector. Since the finest fringe spacing in the interference pattern is related to the spatial coherence of the electron source and thus to the source size, one can deduce an upper limit of the virtual source size of the employed field emitter tip. Experiments carried out by Fink et al. [88] are performed in a low-energy electron point source microscope, as first described in [89]. The emitted divergent electron wave is scattered at a thin gold film placed several hundred nanometers distant from the tip and the interference pattern is recorded in the far-field. The principle setup for the experiments performed by Spence et al. [90] is similar, except that they use an edge of a carbon film as a sample and record the pattern of Fresnel fringes. In both cases, the resolution corresponding to the finest fringe spacing in the interference pattern equals the virtual source size r_v [91], since the angular coherence α is given by the van-Cittert-Cernike theorem to

$$\alpha = \frac{\lambda}{r_\mathrm{v}} . \tag{10.23}$$

A modified setup has been employed by Morin et al. [29, 92]. Here, an additional biprism is implemented, in order to split the initial divergent electron beam in two and create a fringe pattern at the detector. Again, from the fringe distance, one can determine an upper value of the source size.

The sketched experiments yield values for the virtual source size between 7 Å [90] and 1.6 Å [92]. This rather large discrepancy can be explained by the fundamental differences between the applied measuring methods. Besides, it must be taken into account that these values constitute upper limits, because they are affected by adverse effects like vibrations. However, the measured virtual source sizes are in agreement with the rough estimate of several Ångstroms obtained from (10.17).

These findings are relevant to our experiment as well, since the accordant conditions are comparable. Hence, treating the electron source as a point source is justified in the ray tracing simulations carried out in order to characterize the employed micrometer sized electron lens (4.3.3). Besides, when an image of the virtual source is created by means of this electron lens, the size of the focal spot is not limited by the source size but rather by spherical aberrations of the lens (see also Chapter 5). In our setup, the small virtual source size together with the low-aberrations of the electron lens, allows transforming the initially divergent electron beam into a collimated one with a divergence angle of just a few milliradians.

10.2.3 Tip preparation

The field emitters employed in our experimental setup are single-crystalline ⟨111⟩-oriented tungsten wires of 100 µm thickness, electrochemically etched in NaOH. Its typical shape after etching is shown in the SEM images in Fig. 10.7. The radius of curvature at the apex of an etched tip is usually between 20 and 50 nm. The tip is spot welded to a loop mounted on a dedicated holder (see Fig. 10.7 (a)) so that the tip can be resistively heated when implemented in the vacuum chamber in order to clean it and remove atoms adsorbed to the surface. For additional sharpening of etched tips or re-sharpening of tips that got blunt by repeated heating, they can be sputtered in situ with neon. For this, the tip is first heated for several seconds to a temperature of about 2000 K, estimated by visual inspection, in order to obtain a clean but blunt tip. Thereafter, neon of a purity of 99.999% is admitted to the vacuum chamber by a leak valve until a pressure of $2 \cdot 10^{-4}$ mbar is reached. An emission current of 10 µm is subsequently drawn from the tip, by applying a negative

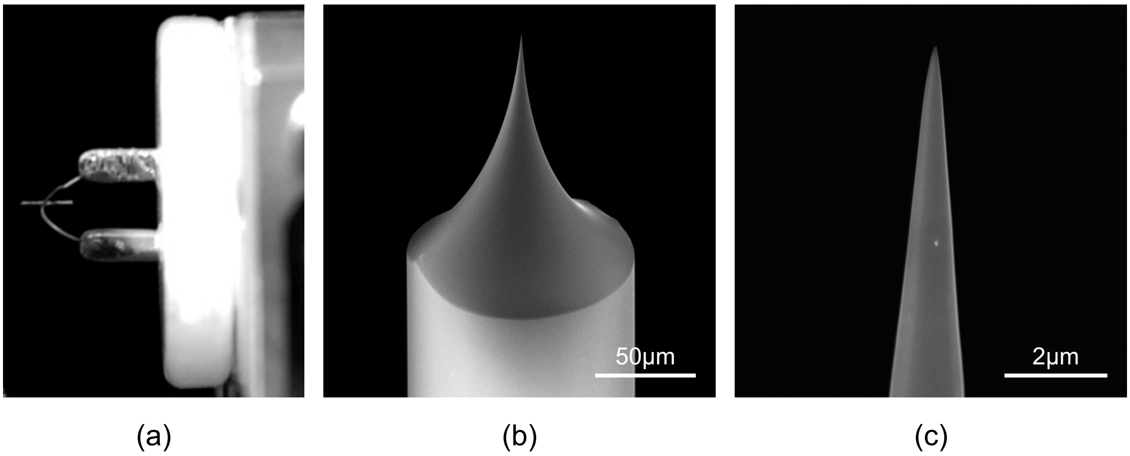


Figure 10.7: *(a) Field emitter tip spot welded to a tungsten loop for resistive heating. (b) SEM image of an etched, sputtered and annealed tip showing its typical shape. (c) SEM image of the same tip at higher magnification.*

voltage between 1.5 and 2 kV. Neon ions are formed by collisions with the field-emitted electrons and are accelerated towards the tip. This leads to sputtering of the tip surface by the impinging ions and to a reduction of the tip radius. Since the power supply is operated in constant current mode, the voltage applied to the tip decreases during the sputtering process, until it stabilizes after a few minutes at a few hundred volts. In order to heal out defects caused by the neon bombardment, the tip is subsequently annealed. This is done at a lower temperature than that required for initial tip cleaning, so that the the tip is not appreciably blunted. The resulting tip radii after this sharpening procedure are around 10 nm. A detailed description of the sputtering process can be found for example in [93].

10.3 Electron lens

As pointed out in Chapter 9, a high spatial coherence and thus a small divergence angle of the incident beam is crucial for obtaining a diffraction pattern corresponding to the Fourier transform of the object function. In principle, a small divergence angle could be achieved by just placing the electron source at a sufficiently large distance and limiting the beam by an additional aperture. Most of the initial beam intensity however would be lost in this way and, because of the limited maximal total current that can be drawn from the field emission tip, the intensities in the higher diffraction orders would become too weak. For these reasons, it is necessary to employ electron optics collimating the initially divergent electron beam. An electrostatic microlens as described in Part I is an appropriate device for this purpose. First, owing to its reduced size, it exhibits reduced spherical aberrations. Thus, when used in combination with a source of an effective size in the Ångstrom range, the achievable divergence angle of the beam is minimized. Second, a significant part of the initial beam current can be gathered by placing the electron source at a distance of only several microns in front of the lens aperture.

The lens consists of two carbon electrodes with micrometer sized circular apertures. The electrodes are separated by an insulating SiN layer of 1 µm thickness. Hence, a voltage of up to 300 V can be applied which results in an electrostatic field distribution around the apertures having a focusing effect on the electron trajectories. When implementing such a microlens in the experimental setup, tip-lens distances are typically between 10 and 20 µm. Usually, the lens is operated in the decelerating mode for diffraction experiments, resulting in electron energies between 30 to 50 eV at the sample. As has been shown in 4.4.1, the achievable half divergence angle of the beam is of several milliradians depending on the tip-lens distance. Since this divergence angle is associated with a finite spatial coherence, this imposes limitations on the maximal size of an object placed at some distance beyond the lens as it will be discussed in detail in 10.5.3.
In the following, the advantages of the implemented microlens in terms of beam intensity will be quantified. Next, the possibility is discussed to adjust the de Broglie wavelength of the electrons by means of reacceleration towards the sample.

10.3.1 Gain of intensity by the implementation of a microlens

With the collimation of the beam its origin is virtually shifted to a larger distance d_v from the aperture (see Fig. 10.8). A half divergence angle θ_v at the lens exit between 3 and 6 mrad, as it has been measured for the microlenses implemented here (see 4.4.1), corresponds to the situation of placing the source at a distance d_v between 50 and 150 µm from a micrometer sized aperture. In order to determine the gain of intensity, the part of the total beam current gathered by the microlens must be compared to the part that would pass the aperture, if the source was placed at a distance d_v. Let θ_t denote the divergence angle of the beam at the lens entrance, as shown in Fig.10.8. The gain of intensity is then given by $\theta_\mathrm{t}^2/\theta_\mathrm{v}^2$. The minimal achievable divergence angle θ_v increases with decreasing tip-lens distance and there is an approximately linear relationship between the tip-lens distance d_t and the distance d_v from the virtual source to the lens as can be inferred from Fig. 10.9 (a). Hence, the same holds for the angles θ_t and θ_v. As a consequence, the gain of intensity is rather independent of the tip-lens distance. The intensity is enhanced by an average factor of about 60 as shown in Fig. 10.9. The values plotted in Fig. 10.9 have been obtained from the measurements of the minimal achievable spot size at the detector (compare section 4.4.1).

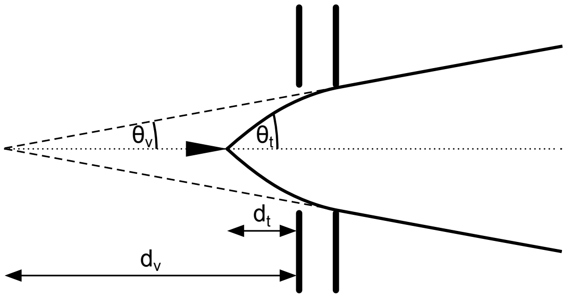


Figure 10.8: The reduction of the beam divergence angle by the microlens leads to a situation equivalent to a source placed at a larger distance d_v from the lens.

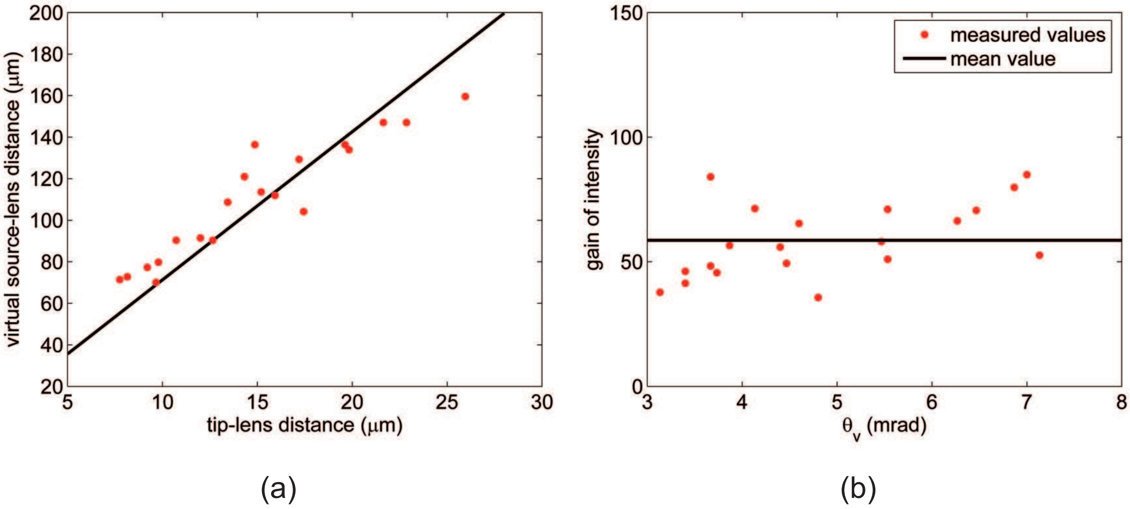


Figure 10.9: *(a) Dependence of the virtual source-lens distance on the tip-lens distance. The values are obtained from the measurements of the minimal achievable spot size at the detector. (b) Gain of intensity due to the implementation of a microlens. According to the measurements of the minimal achievable divergence angles, the intensity can be enhanced by a factor of about 60.*

As a consequence, there is a trade off between the size of the region that can be illuminated coherently on the one hand, and the beam current at the sample on the other hand: reducing the tip-lens distance in order to increase the current, leads to an increase of the divergence angle by the same factor. This in turn reduces the region that can be coherently illuminated at a fixed distance from the lens. Coherent illumination of an object of a certain size can then be ensured by adjusting the lens-sample distance accordingly. This however, leads again to a reduced current at the object. The part of the total current in the collimated beam, that can be used for coherent scattering at the object, depends on the divergence angle of the beam. If the latter is increased by a factor a, the usable current drops by a factor $1/a^2$. Thus, the intensity gain of a^2 due to a reduction of the tip-lens distance by a factor a is cancelled out by the larger lens-sample distance which is required to ensure coherent illumination of the object.

Hence, when performing a coherent diffraction experiment in the described setup, it doesn't make a difference, whether a sufficiently small beam divergence angle at the sample is ensured by increasing the tip-lens distance or by increasing the lens-sample distance. The only way to increase the signal without altering the spatial coherence in the sample plane, is to increase the total emission current at the tip or, of course, the exposure time at the CCD camera.

10.3.2 Post-acceleration towards the sample

Provided that the sample exhibits planar structure (see 10.5.2), the kinetic energy of the electrons can easily be varied by post-acceleration towards the sample. For this, the voltages at the tip and the lens electrodes are lowered together, while the sample and the MCP input are kept at ground potential (Fig.10.10). The latter ensures a field free region behind the sample. Post-acceleration towards the sample introduces an additional degree of freedom with respect to the de Broglie

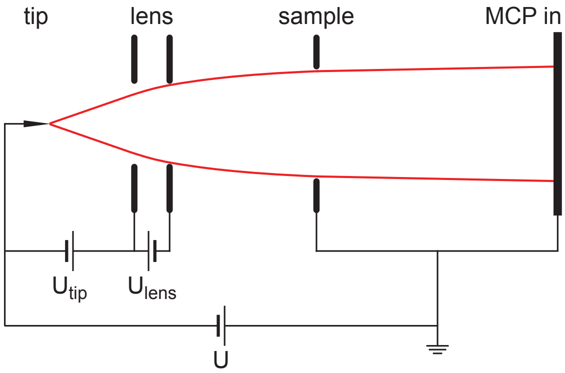


Figure 10.10: *Voltages for post-acceleration of the electrons towards the sample. The potentials at the tip and at both lens electrodes are lowered altogether, resulting in an accelerating field between the second lens electrode and the sample.*

wavelength of the electrons, similar to what could be achieved by implementing a three-electrode lens. While the electron energy at the lens exit is typically around 40 eV, corresponding to a wavelength of 1.9 Å, post-acceleration towards the sample across a potential of up to 250 V yields wavelengths as low as 0.7 Å, enhancing the maximal achievable resolution in the diffraction pattern at the detector. Another advantage is the increase of intensity at the sample due to the reduction of the beam's overall divergence angle. In the case of an homogeneous accelerating field, even spatial coherence would improve, owing to a reduced beam divergence. But unfortunately, such homogeneous field distribution can not be established experimentally due to field penetration at the holes in the sample support film (see Fig. 10.11).

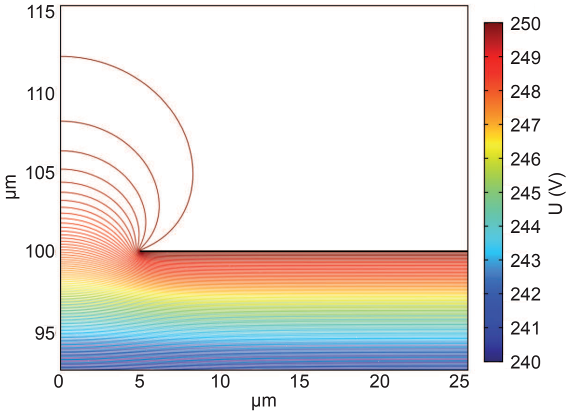


Figure 10.11: *Equipotentials bulging through a 10 µm sized aperture. Equipotentials were calculated for two electrodes separated by a distance of 200 µm and for a potential difference of 250 V. Part of the potential distribution around one of the apertures is shown. The thickness of the electrodes was set to 50 nm. The potential difference between subsequent equipotential lines amounts to 0.1 V. (Calculations were performed with the COMSOL Multiphysics software package [94].)*

This issue has been analysed experimentally in combination with calculations and shall be described in the following. For the experiment, a 1 µm aperture in a SiN membrane coated with platinum on both sides is positioned instead of the lens at a distance of 9 µm from the electron source, limiting the half divergence angle of the beam to $\theta_0 = 56$ mrad. A second coated SiN membrane with a $3 \times 20\,\mu\text{m}^2$ slit aperture replaces the sample and serves as a second electrode. A typical lens-sample distance is $L = 150\,\mu$m. Thus, the second membrane is positioned accordingly. By lowering the potentials at the tip and the first electrode simultaneously, electrons are accelerated towards the grounded second electrode. The kinetic energy of the electrons amounts to $eU_0 = eU_{\text{tip}}$ when entering the accelerating field between the two membranes and increases to the final kinetic energy eU at the second aperture.

Due to geometrical reasons the slit aperture is not completely filled by the beam even when there is no acceleration voltage applied (see Fig. 10.12 (a)). Like this, the change in divergence angle when an accelerating voltage is applied can directly be deduced from the length D in the projection image at the detector as shown in Fig.10.12. The measured reduction of the divergence angle θ_L

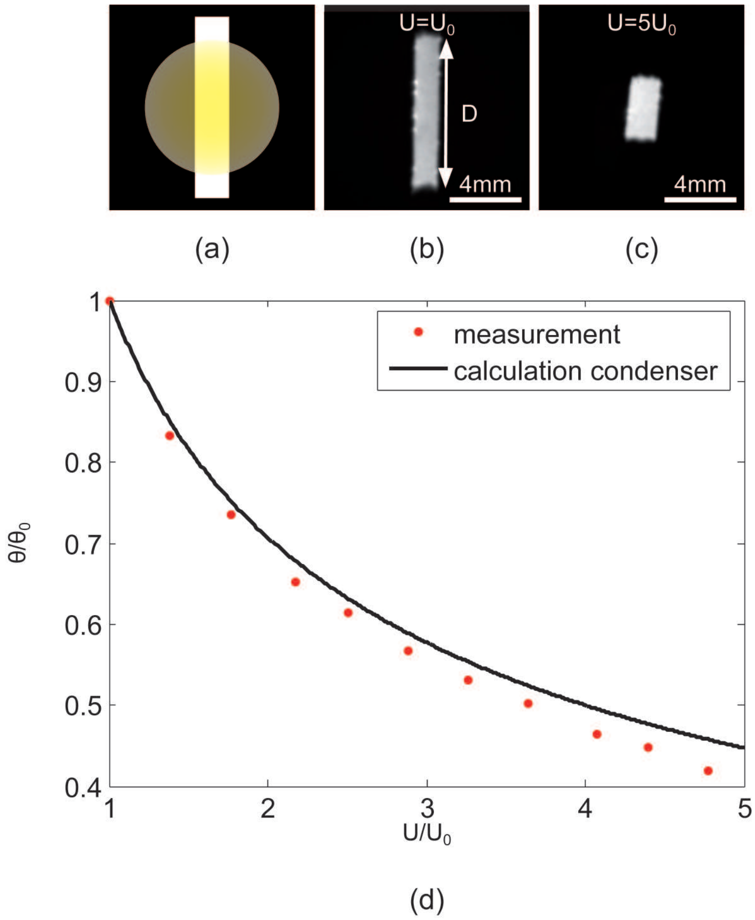


Figure 10.12: *Decrease of the total beam divergence angle by post-acceleration. The electrons are accelerated towards a planar electrode with a rectangular aperture, not completely filled by the beam in vertical direction as depicted in (a). The respective divergence angle can directly be deduced from the lengths D in the projection images as they are shown in (b) and (c) for two different values of U/U_0. (d) The decrease of the overall divergence angle with increasing U/U_0 is in good agreement with the values calculated for a homogeneous accelerating field.*

with respect to the initial divergence angle θ_0 is plotted in Fig.10.12 (d) against the voltage ratio U/U_0 and compared to values calculated for a homogeneous accelerating field. Using the small angle approximation, one obtains

$$\frac{\theta_L}{\theta_0} \approx \sqrt{\frac{U_0}{U}}\,. \tag{10.24}$$

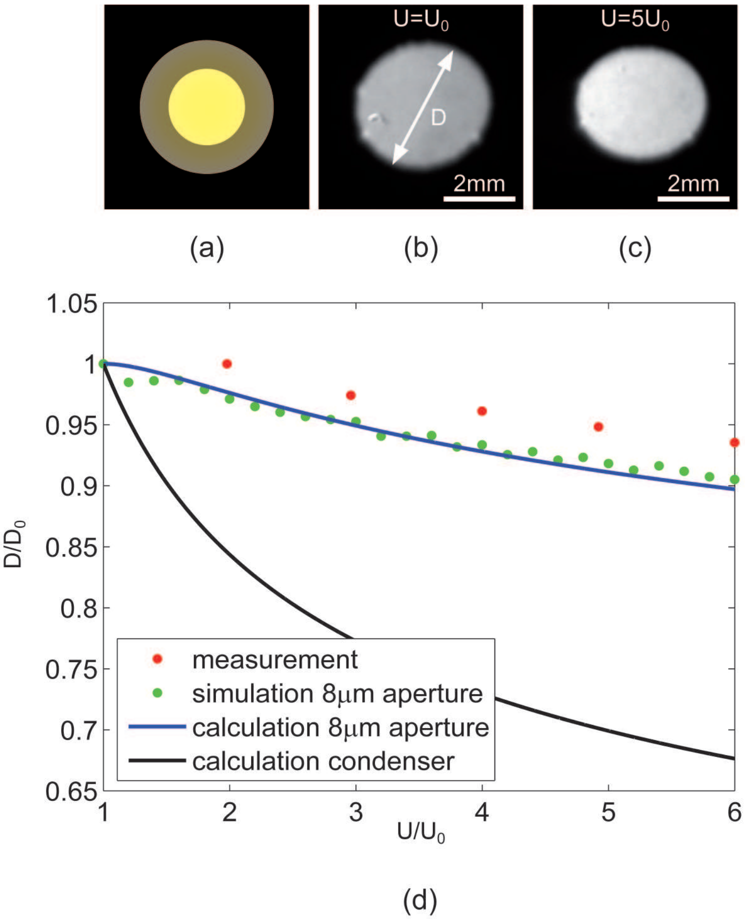


Figure 10.13: *Effect of post-acceleration of the electron beam towards a planar electrode with a round aperture on the size of the aperture's projection image. The aperture with a diameter of 8 µm is completely filled by the beam (a). The size D of the projection image decreases only very slightly with increasing accelerating voltage as can be seen in (b) and (c). The measured values are compared to values obtained from ray tracing simulations, to values calculated for paraxial trajectories passing a homogeneous field followed by an aperture and to values calculated for a homogeneous electric field (d).*

Calculation and measurement for the decrease of the divergence angle θ_L are in good agreement. However, as can be seen in the displayed images, the width of the slit in the projection image does not decrease with higher accelerating voltage as expected for a homogeneous field; it even increases slightly instead. This is due to the penetration of the electrical field through the rectangular aperture, having a much stronger effect in lateral than in longitudinal direction.

In order to improve the comparability between the measured effect of field penetration on the one hand, and calculations or ray tracing simulations on the other hand, an analogue experiment has been performed with a round aperture of 8 µm in diameter in the second membrane, so that the beam diameter at the second electrode is larger than the one of the aperture (Fig.10.13 (a)). As can be seen in Fig.10.13, the decrease of the diameter D of the projection image with increasing accelerating voltage is only very small.

The variation of the size of the projection image, expected for a homogeneous accelerating field, can be calculated using (10.24), taking into account that the initial maximal divergence angle θ_0 now depends on the applied voltage and on the size of the second aperture. This is because with increasing accelerating voltage electrons with higher initial divergence angles θ_0 can also pass the aperture. The initial angle θ_0 of the outermost trajectory passing the second aperture at its radius r_A can be calculated from the following equation for the y-position y_L of a particle accelerated across a homogeneous field of length L [95]

$$y_L = y_0 + \frac{2L}{\sqrt{\frac{U}{U_0}} + 1}\theta_0 \,, \tag{10.25}$$

where y_0 denotes the starting position. This yields

$$\theta_0 = \frac{r_A\left(\sqrt{\frac{U}{U_0}} + 1\right)}{d_t\sqrt{\frac{U}{U_0}} + d_t + 2L} \,, \tag{10.26}$$

approximating the starting position with $y_0 \approx d_t \cdot \theta_0$, where d_t denotes the distance from the source to the 1 µm aperture (see Fig. 10.14 for notation).

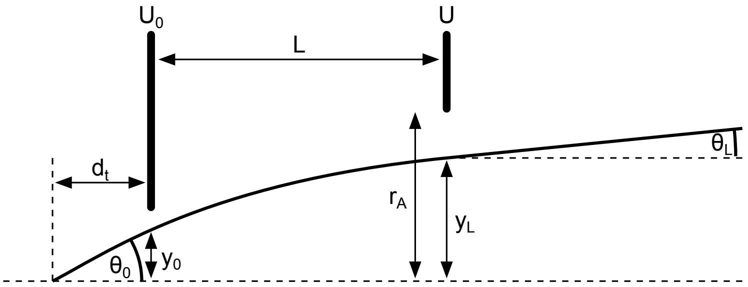


Figure 10.14: *Illustration showing the notation used for the calculation of trajectories in an accelerating field between an aperture at voltage U_0 to a second aperture at voltage U.*

For the case of an homogeneous accelerating field, terminated by a planar electrode with a round aperture, the radial components of the electric field present near the hole lead to an increase in the divergence angle of the electron trajectories when passing the aperture. Considering paraxial trajectories, and thus neglecting an initial component of the particle's velocity in radial direction, the variation of the divergence angle of an electron passing the aperture at a distance y from the optical axis is given by [87]:

$$\Delta\theta = \frac{U - U_0}{4U} \cdot y = -\frac{1}{f} \cdot y \,. \tag{10.27}$$

Using (10.24) and (10.25), it follows that the divergence angle of a trajectory is altered by a uniform accelerating field followed by a round aperture according to:

$$\frac{\theta_L}{\theta_0} = -\frac{1}{f}d_t - \frac{2L}{f\left(\sqrt{\frac{U}{U_0}} + 1\right)} + \sqrt{\frac{U_0}{U}} \,, \tag{10.28}$$

where it has been made use of the fact that the y-position y_L of a particle when passing the aperture is not significantly influenced by the field components in radial direction, and thus is still given by (10.25). The initial divergence angle θ_0 for the trajectory just passing the edge of the aperture is given by (10.26), which can be substituted in (10.28), in order to obtain a value for the size of the second aperture's projection image in dependence on the applied accelerating voltage. These values are displayed in Fig. 10.13 (d) together with the values calculated for the case of a homogeneous accelerating field according to (10.24), the values obtained from ray tracing simulations and the values obtained from the experiments. The experimental result, that the size of the projection image does not decrease significantly, even when the accelerating voltage is such that $U/U_0 = 6$, is confirmed by ray tracing simulations and by the calculations accounting for the effect of an aperture terminating the homogeneous field using the paraxial approximation. In contrast, the decrease of the size of the projection image, that would be predicted for a condenser field, is much stronger, showing that the effect of field penetration through the aperture can not be neglected here.

As can be inferred from Fig. 10.15 the variation of the divergence angle takes place around the aperture, a few micrometers in front of the aperture, as well as a few micrometers beyond the aperture. The distance over which the angle is changing becomes smaller with decreasing aperture size, while the magnitude of the variation of the angle does not depend on the aperture size. Thus, one can conclude that, although the apertures in the sample support film are usually smaller than the apertures in the described experiment, their effect on the electron trajectories can not be neglected. Concerning the spatial coherence of the incoming beam at the place of the sample, one can say, that for a round aperture it is expected to be slightly improved by the whole post-acceleration procedure, however, there is less improvement than for a condenser field. Besides, it is not only the trajectories of the incoming collimated beam that undergo a change in the divergence angle by the penetrating field, but also the trajectories of the scattered electrons. To estimate in what respect this affects the finally recorded diffraction pattern, further simulations are required.

The effects of field penetration through a rectangular aperture on the coherence at the sample as well as on the deviation of the trajectories of the diffracted electrons are not yet clear, since its influence is different in lateral and longitudinal direction. The effect of the equipotentials bulging through the apertures in the support film could in general be reduced somewhat by increasing the film thickness to some hundred nanometers. In fact, mounting the sample on the back side of the support film results in reduced influence of any penetrating fields on the scattered electrons, while detection through large scattering angles is still possible.

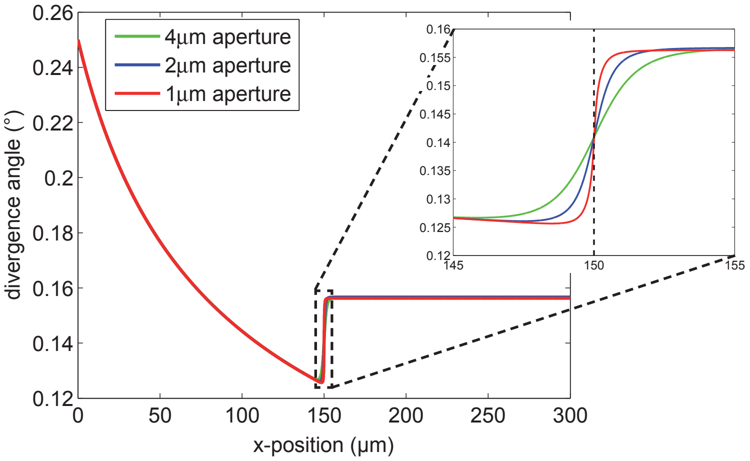


Figure 10.15: *Divergence angle of the trajectory of an electron passing a homogeneous accelerating electric field terminated by a circular aperture placed at a position x = 150 μm. After a reduction of the divergence angle due to the homogeneous field, the angle increases and the trajectory suffers a kink due to the radial components of the electric field around the aperture. (Calculated using SIMION 8.0 [96])*

10.4 The electron detector

The detector is placed at a 75 mm distance from the sample. First, a microchannel plate (MCP) multiplies the incoming electrons. The electrons are then accelerated towards the phosphor coated front side of a fibre optic plate (FOP) kept at a high voltage of 5 - 6kV. The light signal created is then guided to the ambient pressure side, where the image is captured by a CCD camera. Currently we use an MCP with a channel size of 12 μm and a diameter of 45 mm. One pixel in the final image corresponds to 23 μm × 23 μm at the screen. A photograph of the detector flange with the FOP and the objective of the CCD camera is shown in Fig. 10.16. Although the pixel sizes of MCP and CCD chip set a lower limit to the achievable spatial resolution of the detection system, the main limiting factor is the spread of the electron trajectories when passing the 5 mm gap between MCP exit and phosphor screen [97]. This results in a value for the resolution which is considerably larger than the effective pixel size in the recorded images, as will be shown in the following section.

The implementation of a new detector with improved sensitivity, higher spatial resolution, specified with 15 μm, and a diameter of 75 mm is planned. The latter increases the acceptance angle and therewith allows a significantly improved resolution in the diffraction patterns.

10.4.1 Spatial resolution of the detector system

The limited spatial resolution of the detector leads to blurring of the recorded images. This can be described by the convolution of the perfect image, as it would be recorded by an ideal detector, with the non-perfect image of a bright, infinitely small spot, i.e. the image of a delta peak. The latter will be denoted in the following as *detector blur function* and is approximated by a Gaussian function. Describing the image formation at the detector as a convolution is valid if the involved

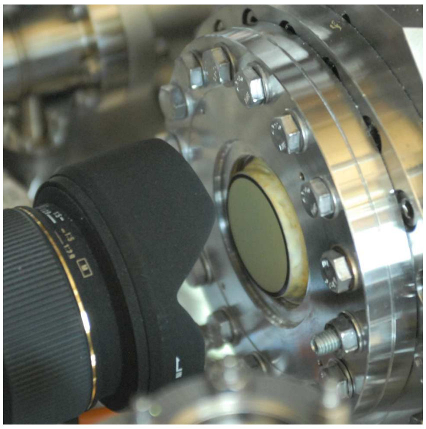


Figure 10.16: *Detector flange with the FOP and the objective of the CCD camera.*

effects are linear. This is true as long as saturation for any of the detector components is not reached.

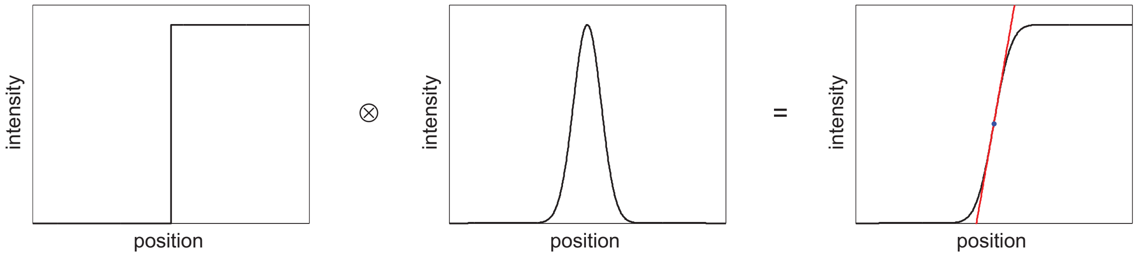


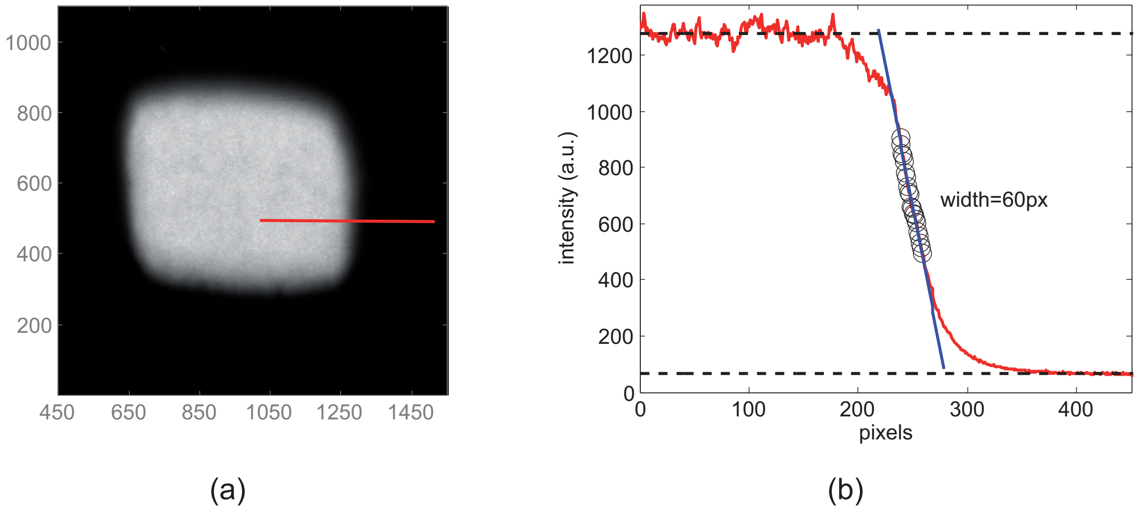
Figure 10.17: *Effect of the detector blur function on the image of a sharp edge. Convolution of a step function with a Gaussian blur function yields a smeared step, described by an error function. From the slope of the tangent to the error function at its inflection point the width of the blur function can be determined.*

For a quantitative evaluation of the detector resolution of our setup, the width of the Gaussian detector blur function is determined. Since a direct measurement is not possible experimentally, it is instead extracted from a projection image of a sharp edge. An intensity profile through the projection image perpendicular to the edge is no longer a perfect step function, but appears smeared out due to convolution with the detector blur function, resulting in an error function, as illustrated in Fig. 10.17.

The full width at half maximum (FWHM) of the Gaussian blur function and the slope c of the tangent to the error function at its inflection point are related via

$$\mathrm{FWHM} = \frac{2}{c}\sqrt{\frac{ln\,2}{\pi}}\,. \tag{10.29}$$

Experimentally, a sharp edge is realized by a rectangular $8 \times 8\,\mu\mathrm{m}$ aperture in a carbon film (see Fig. 10.18 (a)). The distance between electron source and carbon film was chosen to be about $400\,\mu\mathrm{m}$, large enough to prevent significant contributions to the blurring of the edge by diffraction effects and vibrations. To enhance the accuracy of the measurement, the image at the screen is magnified by an objective and recorded by the CCD camera. The pixel size in the final image corresponds to 3.1 µm. Since the width of the step function obtained from an intensity profile through the projection image perpendicular to the edge is large compared to the size of the smallest recognizable features, deteriorations of the image by the optical system can be neglected. In order to determine the width of the step function, intensity profiles perpendicular to the edges in the projection image are evaluated. For noise reduction, this is not done for single line profiles, but profiles along 10 parallel adjacent lines are averaged respectively. Evaluation is then performed as depicted in Fig. 10.18 (b). A linear function is fitted to the ±10 values around the inflection point, yielding the slope of the corresponding normalized and background corrected profile. With this, the full width at half maximum (FWHM) of the detector blur function for the respective settings can be calculated according to (10.29).



(a) (b)

Figure 10.18: *Profile through the projection image of an aperture in a carbon film, that can be regarded as a sharp edge. The width of the step is determined by fitting a linear function to the values around the inflection point. One pixel in the image corresponds to 3.5 µm.*

In order to improve the resolution of the detector the voltage at the screen can be increased as shown in Fig. 10.19, since this reduces the spread of the electrons on their way from the MCP to the screen. Each of the plotted values in Fig. 10.19 is a mean value obtained from the evaluation of eight different profiles through images taken with an emission current of 300 nA, a maximal voltage of 1 kV at the MCP and an acquisition time of the camera of 150 ms. Any influence of other parameters than the screen voltage on the detector resolution could not be observed. Thus, the best achievable resolution of the implemented detector system is about 150 µm. With respect to sampling aspects, this corresponds to an effective pixel size of 75 µm. As it will be deduced in 10.5.3, this does not allow to oversample the diffraction pattern of an object of several hundred nanometers in size.

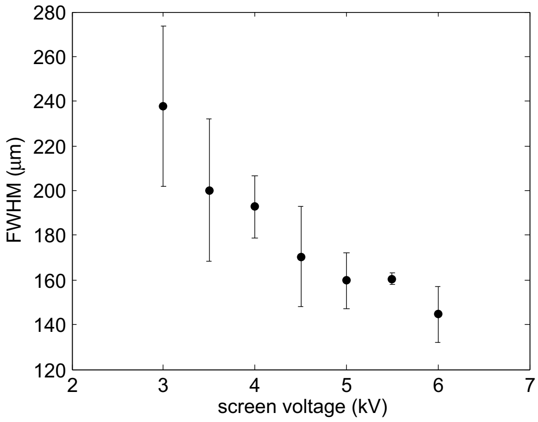


Figure 10.19: *Detector resolution. FWHM of the detector blur function in dependence on the voltage at the screen. Each of the plotted values is a mean value of eight different profiles through projection images of a sharp edge. Images have been recorded with an emission current of 300 nA, maximal intensifying voltage at the MCP of 1 kV and an acquisition time of the camera of 150 ms.*

10.5 Sample preparation

For any kind of coherent diffraction experiment the sample preparation must be such that contributions to the diffraction pattern by a sample mount or support film are minimized. For subsequent numerical phase retrieval, the object size, understood as the region contributing to the diffraction pattern, must not exceed a certain value. The latter is given by the sampling rate at the detector, the spatial coherence of the beam and the desired oversampling ratio. In the following, the region contributing to the diffraction pattern, will also be referred to as *physical support*.

Although the aim of the experiment is imaging individual biological molecules, carbon nanotubes (CNTs) were used as samples for first diffraction experiments with coherent low-energy electrons. Due to their well known structure they appear to be suitable test objects. Carbon nanotubes are commercially available in different forms and have the advantage of being quite robust also with respect to illumination with electrons of higher kinetic energies or high currents and they are relatively easy to handle. Moreover, once established, preparation procedures for carbon nanotubes might be extended for the preparation of proteins as the latter may readily be attached to the nanotubes.

For preliminary experiments we used a method for the mounting of carbon nanotubes, which does not allow the recording of oversampled diffraction patterns, since the physical support can not be limited to a sufficiently small size. This method will be described in 10.5.1. A general discussion about the possibilities for limiting the physical region of support is provided in section 10.5.2. In 10.5.3 the implications of pixel size and beam coherence on the size of the physical support are analysed quantitatively for our experimental setup. Subsequently, I will describe a procedure for the preparation of carbon nanotubes on holey carbon films, offering the possibility to adjust the size of the physical support according to the respective parameters and the possibility to be extended for the preparation of proteins. A brief survey of the established methods for attaching proteins to carbon nanotubes will be given in 10.5.5.

10.5.1 Carbon nanotubes attached to a tip

For preliminary diffraction experiments carbon nanotubes have been attached to the apex of an etched tungsten tip. This allows to place the carbon nanotube in the parallel beam right behind the lens. A schematic drawing to this is shown in Fig. 10.20. The procedure to attach a nanotube to the apex is the following: carbon nanotubes are spread over a holey carbon film from a dispersion with a concentration such that the nanotubes form a network all over the carbon film and also across the holes. The sample is inspected in the projection mode of the experimental setup, i.e. it is illuminated with a divergent electron beam. Very fine structures on the sample, such as individual nanotubes, can be identified due to the interference fringes appearing on the screen, when the sample is approached by the field emitter tip down to a distance of a few micrometers. Such a hologram of a very fine structure within a nanotube network is displayed in Fig. 10.21. In order to pick up an individual nanotube, the apex of the sample tip first is brought in physical contact with the nanotube. Next, several milliamperes are sent through the tube in order to weld it to the tip apex and remove it from the network. A projection image of nanotubes successfully attached to the sample-tip is shown in Fig. 10.21 (b).

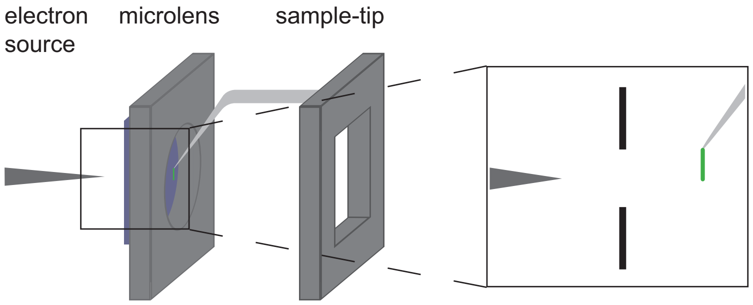


Figure 10.20: *Schematic drawing of the geometrical arrangement of sample and lens, when a carbon nanotube is attached to a tip. The sample-tip is mounted on the movable sample holder such that it can be positioned in the parallel electron beam and manoeuvred closely behind the lens.*

With respect to diffraction experiments, this sample preparation method results in a physical support being limited by the beam diameter, i.e. being larger than 1 µm. Taking into account the achievable beam divergence angle and the given detector properties, it is however not possible to record an oversampled diffraction pattern of a nanotube attached to a tip. This is because the beam diameter is too large to fulfil the conditions for oversampling and spatial coherence as will be discussed in more detail in 10.5.3. Nonetheless, for the very first diffraction patterns of individual carbon nanotubes obtained with our setup, the samples have been prepared in the way described. The patterns are presented in 11.2.1.

10.5.2 Limiting the physical region of support

In the idealized case of a free-standing object, the region contributing to the diffraction pattern is either limited by the object itself, as it is illustrated in Fig. 10.22 (a), or by the diameter of

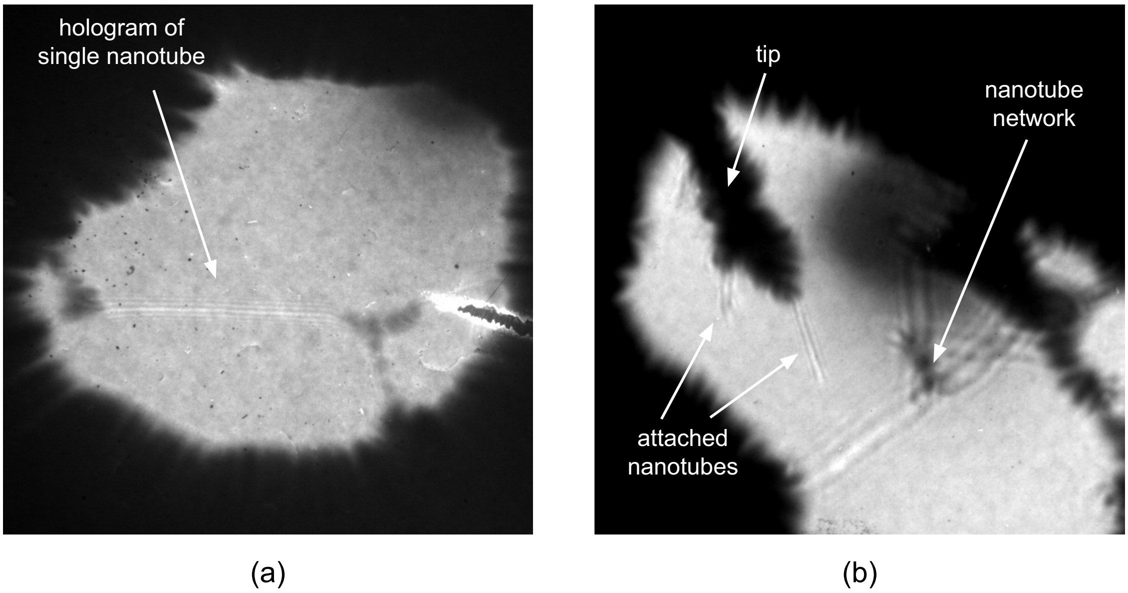
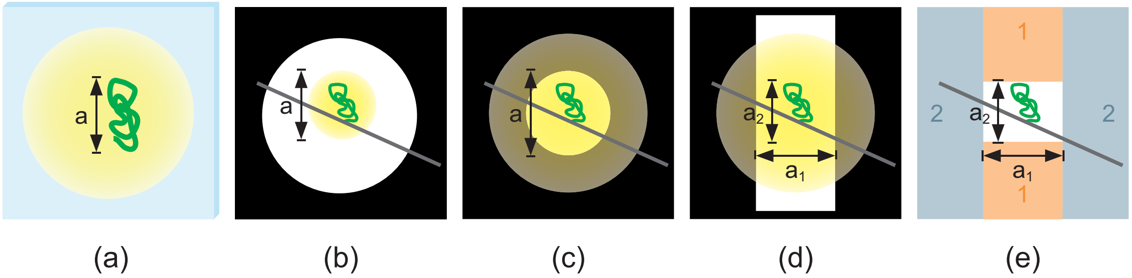


Figure 10.21: *Attaching a nanotube to the apex of an etched tungsten tip. (a) Hologram of a nanotube within a nanotube network prepared on a holey carbon film. (b) Nanotubes picked up with the sample-tip from a nanotube network.*

the illuminating beam (Fig. 10.22 (b)). Such situation can be realized at least to some degree by placing the sample on a transparent substrate not showing any structure on the scale of the strived resolution.

For a coherent diffraction experiment with visible light for example, the sample can be placed between two clean glass slides causing only a constant phase shift and thus not contributing to the recorded diffraction pattern [98]. In the case of x-rays, a thin membrane, e.g. a silicon nitride membrane of several hundred nanometers in thickness, is an appropriate substrate [42, 74], at least for imaging at resolutions down to several nanometers. For the experiments that are planned at the XFELs for coherent diffraction imaging of biomolecules methods for container free sample handling, e.g. spraying techniques, have been proposed [7]. These indeed, would allow to illuminate the molecules in free space. For high-energy electrons, with energies conventional for TEM, amorphous carbon films constitute a transparent substrate and are therefore widely used as sample support films. However, the irregular structure of these films results in a background signal in high-energy electron diffraction patterns, attenuating the signal of supported particles and thus limiting the achievable resolution. Therefore, it has been suggested to use very thin crystalline sheets as a support instead [99, 100], such as thin graphite layers or even graphene or graphene oxide. These layers are transparent for high-energy electrons and do not show any structure at resolutions down to the periodicity of the lattice. Moreover, at higher resolutions the periodic signal resulting from the crystallinity of the layers could easily be subtracted from the diffraction images [101]. When free-standing graphene layers [100, 102] as well as graphene-oxide layers [103, 104] became available a few years ago, their suitability as support films for TEM sample preparation was examined and first images of biological molecules on graphene oxide films have already been reported [101, 105, 106]. However, whether routine preparation of biological molecules on graphene or graphene oxide will ever be possible, remains to be seen. While amorphous carbon films of 10 nm in thickness as used for TEM sample preparation are known to be opaque for low-energy electrons [107], it is no yet clear whether graphene or graphene oxide is sufficiently transparent for low-energy electrons with energies between 50 and 250 eV used in our coherent diffraction experiments. First measurements on the transmission of low-energy electrons with energies between 2 eV and 317 eV through graphene have been performed [107], indicating that there is some transmission through free-standing graphene layers. However, it is not yet clear which part of the measured transmitted signal can be attributed to unscattered electrons and which part is due to incoherent secondary electrons. Preliminary examinations have also been performed in our experimental setup, confirming these results qualitatively. In addition, these experiments revealed that cleanliness of the graphene films over areas of at least some hundreds of square nanometers is difficult to achieve.



(a) (b) (c) (d) (e)

Figure 10.22: *Different possibilities for experimentally limiting the region contributing to the diffraction pattern. (a) Object limiting the physical region of support. (b)The physical support is limited by the beam diameter. (c) The size of the physical support is determined by an aperture in a non-transparent support film. (d) The sample is suspended over a rectangular slit, the region of support then being limited by the support film in one direction and by the object size in the other direction. This leads to a more complicated support constraint in the reconstruction algorithm, since part of the region outside the support is known to be completely transparent (region 1), while the other part is known to be opaque (region 2), as illustrated in (e).*

In the experiments on coherent diffraction with high-energy electrons conducted so far, where different types of carbon nanotubes have been imaged, the samples were suspended over microstructures so that they were at least partly free-standing [45, 108]. The region contributing to the diffraction pattern was limited in these cases by the beam diameter to 50 - 100 nm (see also Fig. 10.22 (b)).

Suspending the molecules of interest over holes or other microstructures in thin conductive films appears to be a suitable method also for coherent diffraction microscopy with low-energy electrons. Not any protein, however, can be suspended over a hole in a thin film because of shape and chemical properties. Instead, carbon nanotubes shall serve as a template to which the proteins can be attached by biochemical means. This template-protein complex can then be suspended over holes in a support film. In contrast to the diffraction experiments carried out with high-energy electrons in TEMs, where the beam can be collimated to a spot of 50 nm in diameter in the object plane, the beam diameter in our setup is larger than one micrometer. Hence, as mentioned before, oversampling of the diffraction pattern with the beam diameter limiting the physical region of support is not possible. Instead, the physical region of support must be limited to an appropriate size by an aperture, cutting off the illuminating beam. This can be achieved by suspending the template-protein complex over holes of only a few hundred nanometers in diameter in a non-transparent sample support film, as illustrated in Fig. 10.22 (c).

The diffraction pattern resulting from such a structure is a superposition of the diffraction pattern of the sample, i.e. the nanotube with the attached protein, and the diffraction pattern of the aperture in the support film. Provided that the diffraction pattern is recorded with an appropriate sampling rate, the phase retrieval algorithm will lead to a reconstruction of the object together with the masking aperture. One can make use of the knowledge about the shape and size of the aperture by putting it in the iterative algorithm as support constraint. As shape and size are known quite precisely, a tight support constraint can thus be imposed, which is expected to speed up the algorithm's convergence [68].

On the other hand, the superposition of the diffraction pattern of the aperture may be disturbing during the experiment, as it makes visual recognition of the object's diffraction signal more difficult. Thus, a sample suspended over a rectangular aperture, as illustrated schematically in Fig 10.22 (d), would be advantageous from the experimental point of view, because the signal in the diffraction pattern of the nanotube perpendicular to the tube direction is not overlaid by the signal due to diffraction at the aperture edges. For the reconstruction again, this would result in more complicated support constraints. The region of support is limited in one direction by the aperture, where the region outside is known to be opaque (region 2 in Fig. 10.22 (d)). In the other direction, a region of support can be defined according to the approximate size of the object, where the area outside is known to be transparent (region 1 in Fig. 10.22 (d)). Reconstruction algorithms adapted to such a varied support constraint are conceivable but have not yet been implemented.

However, the preparation of carbon nanotubes over rectangular apertures is of advantage mainly for preliminary experiments, aiming at establishing the experimental conditions for obtaining diffraction patterns of individual carbon nanotubes rather than at recording of oversampled diffraction patterns. Once sample preparation techniques as well as the recording of diffraction patterns are a matter of routine, it seems to be the better choice to limit the region of support by a small circular aperture in order to record reconstructable oversampled diffraction patterns. A quantitative discussion about the implications of the detector pixel size and the spatial coherence of the incident electrons on the size of the physical support shall be provided in the following.

10.5.3 Size of the physical region of support

The maximal allowed size a of the region contributing to the diffraction pattern is limited on the one hand by the sampling condition to

$$Oa = \frac{\lambda R}{\Delta y}, \tag{10.30}$$

where R denotes the distance to the detector, and Δy is the sampling rate. Since the oversampling ratio O must be larger than 2 (see Chapter 8), an illuminated region of 1 μm in size would require an unrealistically small pixel size of the detector. For the current effective pixel size of $\Delta y = 75\,\mu\text{m}$ (see section 10.4), the size of the physical support should be limited to less than 50 nm for a wavelength of $\lambda = 1\,\text{Å}$ to ensure an oversampling ratio of $O \geq 2$. For the detector planned to be implemented in the setup the effective pixel size is specified with 15 µm resulting in a maximal value for Oa of 500 nm for $\lambda = 1\,\text{Å}$.

On the other hand, the size of the illuminated region is limited by the spatial coherence, i.e. the finite divergence angle θ_v of the illuminating beam, according to

$$Oa \leq \frac{\lambda}{2\theta_\text{v}} \,. \tag{10.31}$$

as has been deduced in 9.1. The half divergence angle can be associated with the distance d_v from a virtual source to the sample, defined by

$$\theta_\text{v} = \frac{Oa}{2d_\text{v}} \,. \tag{10.32}$$

Thus, the relation between the virtual source distance and the maximal coherently illuminated region is

$$Oa = \sqrt{d_\text{v}\lambda} \,. \tag{10.33}$$

Typical values for the distance from the virtual electron source to the lens are in the range of 100 µm, as it has been deduced in 10.3.1. Thus, values for the virtual source-sample distance between 200 and 600 µm are experimentally realistic with regard to sample positioning issues and to the intensity in the diffraction pattern. The dependence of the largest possible coherently illuminated region Oa on the virtual source-sample distance is plotted for two different wavelengths in Fig. 10.23 and compared to the limitations imposed on Oa by the detector pixel size of 75 µm for the detector currently implemented. It becomes obvious that the limitations imposed on Oa by (10.30) and (10.33) for a pixel size of 75 µm are in the same range and it depends on the exact values of the several parameters, which of the two conditions is the limiting one. For an effective pixel size of 15 µm, as specified for the planned detector, the stricter limitation will be imposed by the spatial coherence, while in the case of the current detector it is rather the spatial resolution, imposing the stricter limitation on the object size. For the actual effective pixel size of 75 µm, the size of Oa, i.e. the physical region of support multiplied with the oversampling ratio, is required to be smaller than 100 - 200 nm. This means, that the specimen should be prepared over a hole of at most 100 nm in diameter, in order to achieve an oversampling ration of $O = 2$. This however seems unrealistically small from the experimental point of view, namely concerning the inspection of the sample when it is placed beyond the lens and with respect to positioning issues.

10.5.4 Preparation of carbon nanotubes over holes in a support film

In the following, the procedure for the preparation of carbon nanotubes suspended over holes or slits in a support film shall be described. While, in a first step, the nanotubes themselves serve as sample molecules, the method described here could also be modified for the preparation of

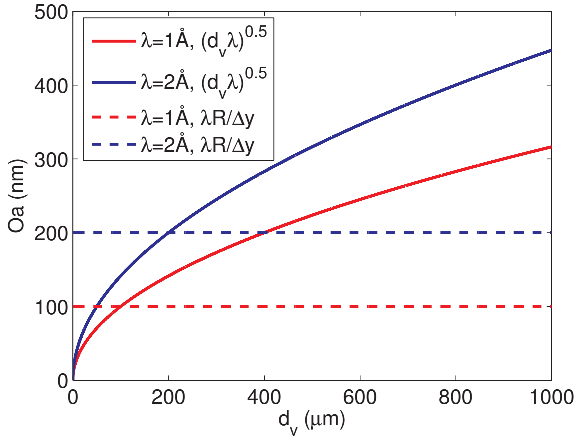


Figure 10.23: *Limitations imposed on the size of the region contributing to the diffraction pattern by spatial coherence and the detector pixel size. The dependence of the maximal value of Oa on the virtual source-sample distance* d_v *and the limitations imposed on Oa by the detector pixel size of* $\Delta y = 75\,\mu m$ *are shown in the graph for wavelengths of 1 Å and 2 Å.*

samples with proteins suspended over holes using carbon nanotubes as templates.

Commercially available carbon films on standard TEM grids are used as support films. Carbon is a suitable support material, due to its relatively low secondary electron yield and due to low variations in surface potentials. Besides, using standard support films, allows to profit from established TEM preparation methods. The carbon support films can be structured with holes or slits of the desired size using a focused ion beam.

The methods for getting the nanotubes lying over holes in the carbon film depend on what kind of nanotubes is used, i.e. if the nanotubes are dispersed in water or in an organic solvent.

Preparation of functionalized carbon nanotubes

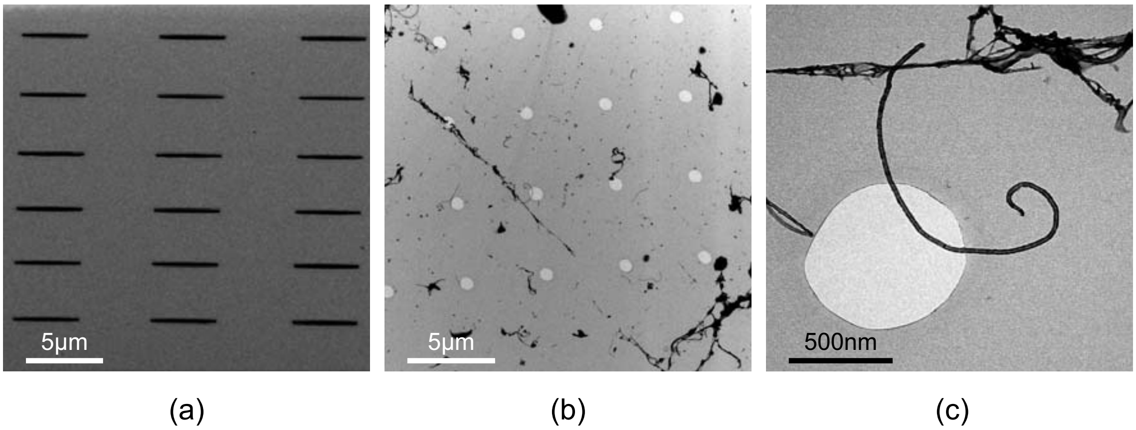
The preparation method described in the following can be applied for the preparation of carbon nanotubes exhibiting a hydrophilized surface, i.e. they can be dispersed in water. The nanotube dispersions used for the work presented here, have been purchased from NanoLab [1]. Samples with bamboo structure multi-walled carbon nanotubes have been prepared from a dispersion with a concentration of 40 mg/l, the suspension of double-walled carbon nanotubes exhibits a concentration of 0.5 mg/l.

For hydrophilization the carbon support film is treated in an UV-ozone cleaner for 30 minutes. In the meantime, the nanotube dispersion is sonicated with an ultrasonic horn dipped directly in the liquid for 2 minutes, in order to separate the nanotubes. A droplet of 2 µm of the dispersion is put onto the prepared carbon foil. Due to the hydrophilic surface of the carbon film, the droplet spreads out yielding a uniformly thin water film. The sample is then vitrified in liquid nitrogen. This must be done very quickly, to prevent binding of the nanotubes to the carbon film. After sublimating the layer of amorphous ice by freeze-drying the sample at −100°C, it is warmed up again to room

---

[1] www.nano-lab.com

temperature in vacuum. It has turned out that, compared to air drying, the yield of nanotubes lying over holes is much higher when the freeze-drying method is applied [75]. This is because, during the air drying process, there is time enough for the nanotubes to find their energetically optimal position, and they end up bound to the carbon film. If the droplet of nanotube suspension is shock-frozen, in contrast, the nanotubes are kept in their momentary position and just "fall" down on the carbon film when the surrounding ice is sublimated.
A TEM image of bamboo structure multi-walled carbon nanotubes prepared in the described manner is shown in Fig. 10.24. However, it has turned out, that if double-walled carbon nanotubes are prepared accordingly, the structures lying over the holes in the support film are mostly bundles of several nanotubes, while individual nanotubes are rarely found. The consequences on the diffraction patterns obtained from these structures will be discussed in Chapter 11.



(a) (b) (c)

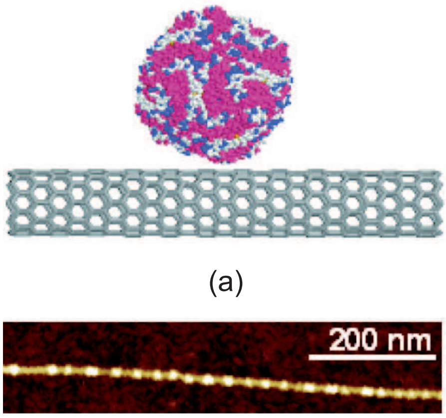
Figure 10.24: *(a) SEM image of a carbon support film structured with slits by ion milling. (b) and (c) TEM images of multi-walled CNTs prepared by the described procedure on a holey carbon support film.*

Preparation of carbon nanotubes dispersed in an organic solvent

When other solvents than H_2O are used for dispersing the carbon nanotubes, freeze-drying is not an appropriate method. For nanotubes dispersed in organic solvents, as for example dimethylformamide or chloroform, a less elaborate preparation method is used. The carbon foil is again first cleaned in UV-ozone and then heated to 180°C on a stove. After sonicating the nanotube dispersion for 2 minutes a droplet of 2 µm is put on the warmed film. The solvent evaporates within a few seconds, leaving some of the nanotubes stretched over holes. This preparation method is quite fast, however it can not be applied for the preparation of any biological molecules attached to the nanotubes.

10.5.5 Carbon nanotubes as templates

There are several established techniques for attaching proteins to the surface of carbon nanotubes that can be found in literature. One possibility is the binding of the proteins to the nanotube surface by nonspecific adsorption. This occurs, when as-grown nanotubes are just put into a protein solution and can be attributed to hydrophobic interactions between the hydrophobic regions of the protein and the hydrophobic surface of the nanotube [109]. Nonspecific binding to carbon nanotubes has been observed for a variety of mainly globular proteins, among them streptavidin, avidin, bovine serum albumin, α-glucosidase, staphylococcal protein A [109, 110, 111]. An illustration of globular protein adsorption to a nanotube is shown in Fig. 10.25 together with an AFM image of a nanotube decorated with non-specifically adsorbed protein A.



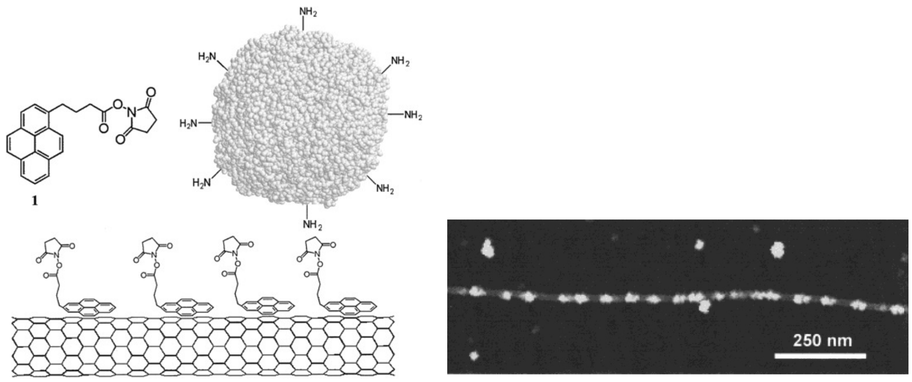
(a)

(b)

Figure 10.25: *Non-specific binding of proteins to carbon nanotubes. (a) Proteins tend to adsorb non-specifically to as-grown nanotubes due to hydrophobic interactions. (b) An AFM image of a staphylococcal protein A adsorbed non-specifically to the surface of a nanotube. (Images taken from [109])*

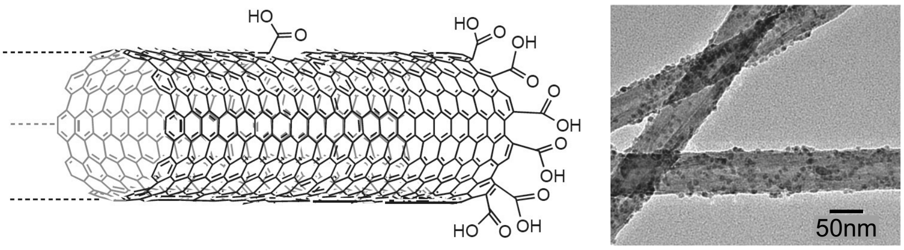
Another approach for the binding of proteins to the walls of a nanotube is based on the idea to use a relatively small molecule as a linker, containing a 'tail' interacting with the protein and an aromatic ring that irreversibly adsorbs to the nanotube surface by π-stacking forces. Since this is the same interaction as present between the layers of graphite, attachment of a protein in this way has the advantage to preserve the sp^2 nanotube structure and thus their electronic characteristics. This method of non-covalent functionalization of nanotubes was developed in Hongjie Dai's laboratory [112] and is illustrated in Fig. 10.26. A molecule is used possessing a planar pyrenyl group to bind to the nanotube and a 'tail' consisting of a succinimidyl ester group (molecule 1 in Fig. 10.26). The latter reacts with amine groups, present at the surface of most proteins, to form an amide bond.

Besides, there are several procedures to covalently attach proteins to carbon nanotubes, making use of COOH groups at the surface of the nanotube. Functionalization of the nanotube surface with carboxyl groups is achieved by acid treating or other oxidizing methods like ozonolyis or air oxidation. As illustrated in Fig 10.27, carboxyl groups are formed at the ends of the nanotube and at other defect sites, created by the oxidizing treatment. The proteins can then be attached to the nanotubes applying a standard method for binding biomolecules to other materials: The carboxyl groups are first converted to active esters via diimide-activation, and then the active esters are reacted with the amine groups on proteins resulting in the formation of an amide bond between the carbon nanotubes and the protein. Successful applications of this method for attaching different proteins to carbon nanotubes are reported for example in [113, 114, 115].



(a) (b)

Figure 10.26: *Non-covalent binding ot proteins to CNTs. (a) Amine groups on a protein react with succinimidyl ester anchored to the nanotube surface via a pyrenyl group. (b) AFM image showing a bundle of single-walled CNTs functionalized with ferritin molecules by non-covalent binding. (Images taken from [112])*



(a) (b)

Figure 10.27: *Covalent functionalization (a) Model of a carbon nanotube functionalized with carboxyl groups at defect sites (from [116]). (b) TEM image of ferritin attached to a multi-walled nanotube by covalent binding (from [115]).*

Chapter 11

Diffraction patterns of carbon nanotubes

First experiments on coherent diffraction with low-energy electrons have been performed using carbon nanotubes as test-objects. Carbon nanotube samples have been prepared according to the procedures described in 10.5. After a brief introduction about the intrinsic characteristics of a diffraction pattern of a carbon nanotube, first experimental diffraction patterns obtained in our setup will be shown in the following. Some issues arising from the recording of experimental diffraction patterns and possibilities to tackle them by post processing of the images will be discussed along with the presented diffraction patterns.

11.1 Calculated diffraction pattern of a carbon nanotube

In order to get an idea what a diffraction pattern of a carbon nanotube should look like and what we expect to to see at the detector, diffraction patterns can be calculated from molecular models of the nanotubes applying the approaches described in 7.2. The simplest way to obtain an estimate of the diffraction pattern of a carbon nanotube is by applying the approximation described in 7.2.1. It is based on the assumption that the curvature of the Ewald sphere in reciprocal space can be neglected and hence the diffraction pattern corresponds to the two-dimensional Fourier transform of the projected scattering potentials. This can easily be implemented by projecting a molecular model of a carbon nanotube, the atoms being represented as spheres, onto a plane parallel to the tube axis and calculating the Fourier transform of the resulting two-dimensional image. Subsequently the image is smoothed with a Gaussian filter, which means that the atomic potentials are represented by Gaussian functions of arbitrary width. This approximation, however, is only valid for small scattering angles or for two-dimensional objects. Hence, for tube diameters of several nanometers and a maximal scattering angle of 17° determined by the detector size, an experimental diffraction pattern is expected to deviate from the calculated one. It is difficult to predict in what respect the diffraction pattern would be affected by these deviations since this strongly depends on the tube diameter, its atomic structure and also on the wavelength of the incident radiation. The most important features can nonetheless be extracted from a diffraction pattern calculated using the

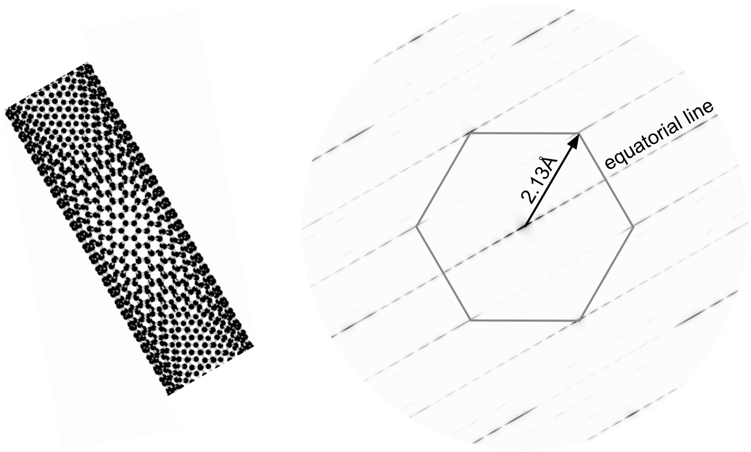


Figure 11.1: *Molecular ball-and-stick-model of a zigzag (20,0) carbon nanotube and its diffraction pattern calculated using the projection approximation. The main features of the diffraction pattern are the equatorial line perpendicular to the tube axis with modulations corresponding to the tube diameter and the hexagonal symmetry of the intensity peaks.*

said approximation as it is shown in Fig. 11.1 for a single walled carbon nanotube. More exact simulations could be obtained by applying the Born approximation described in 7.2 to a three-dimensional molecular model of a carbon nanotube. This means, its three-dimensional Fourier transform is calculated and the values on the Ewald sphere are subsequently extracted.

The most striking feature of the diffraction pattern shown in Fig. 11.1 is the equatorial line, perpendicular to the tube axis with intensity modulations associated to the tube diameter. Since the Fourier transform of a rectangular function of width a is proportional to $\operatorname{sinc}(\pi a u)$, the distance between the intensity maxima along the equatorial line is given by the reciprocal of the tube diameter. In experimental diffraction patterns, the distance between the modulations may deviate somewhat from this value, especially for the higher order maxima. The cause is the aforementioned Ewald sphere curvature. Depending on the exact structure of the tube, it might also happen, that some of the modulation maxima are missed by measuring on a spherical surface in reciprocal space instead of on a plane through the origin [117].

The symmetry of the carbon lattice is represented in the diffraction pattern by the hexagonal arrangement of the intensity peaks. For a nanotube with arbitrary rolling angle, the orientation of the carbon lattice at the top side of the tube is different from the one at the bottom side. Thus, two sets of hexagonally arranged peaks are superposed in the diffraction pattern. The diffraction pattern shown in Fig. 11.1 was calculated from the model of a single-walled zigzag nanotube with the indices (20,0), where the top and the bottom side of the tube exhibit the same lattice orientation, therefore yielding only one set of peaks. The diffraction patterns of a double- or multi-walled carbon nanotube becomes more complicated, since several lattice structures of different orientation and structures of different diameters are contributing. The distance of the hexagonally arranged peaks from the zero-order spot is given by multiples of the length of the reciprocal lattice vector corresponding to a frequency $u = (2.13\,\text{Å})^{-1}$. This means that even with a wavelength of 0.7 Å the diameter of the detector currently implemented in our experimental setup, does not suffice to

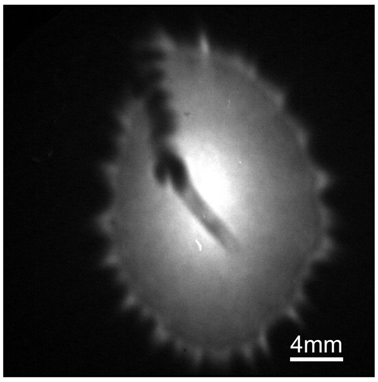


Figure 11.2: *Projection image of a carbon nanotube attached to the tip positioned behind the 1 μm aperture of the lens.*

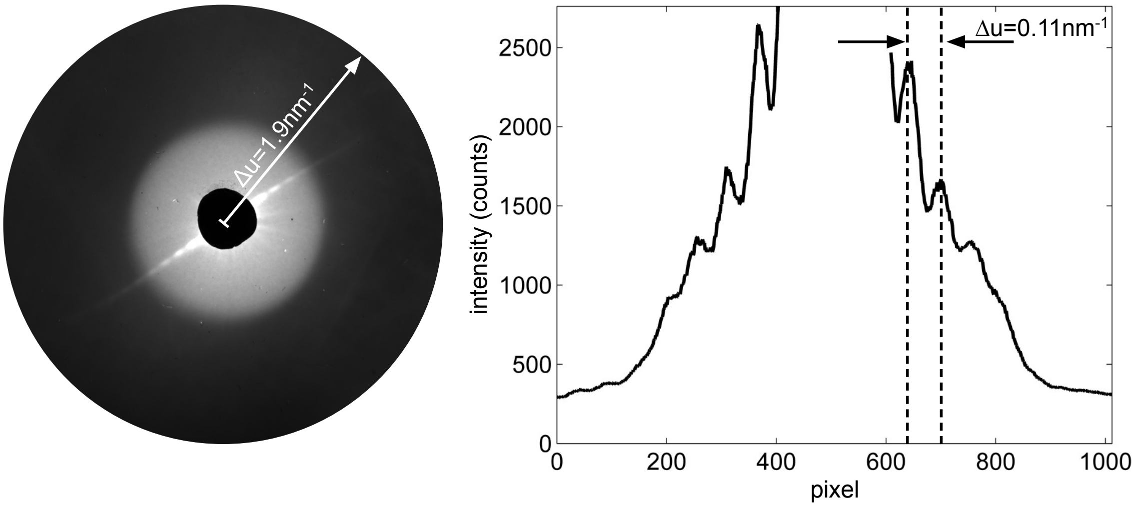
detect these spots. Thus, only the modulations of the equatorial line are expected to be present in the recorded images.

The design of the planned detector in contrast, with its 75 mm in diameter, would allow for a sufficiently high resolution. For experimental detection of the peaks corresponding to the carbon lattice it is, however, crucial that the part of the carbon nanotube exposed to the beam is straight and stable. If it is curved or vibrating, the one-dimensional periodicity of the tube is lost, resulting in smeared features of the diffraction pattern. This has also been pointed out by Meyer et al. [108] who performed diffraction experiments at single-walled carbon nanotubes with high-energy electrons. It shall be stressed here, that the main advantage of coherent diffraction imaging just consists in its applicability also to non-periodic structures. The information about the overall shape of a nanotube as well as about its structure on an atomic scale are contained within the continuous speckles of the diffraction pattern and can be retrieved when this diffraction pattern is oversampled. Hence, an enhanced spatial resolution of the detector allowing to record oversampled diffraction patterns, would facilitate sample preparation in the sense that the structure of a carbon nanotube could be revealed from its diffraction pattern independent of the overall shape of the tube.

11.2 Low-energy electron diffraction on carbon nanotubes

11.2.1 Experimental diffraction patterns of carbon nanotubes attached to a tungsten tip

A multi-walled carbon nanotube is attached to a tungsten tip as described in 10.5.1. In the projection mode of the microscope the sample-tip is positioned beyond the lens aperture and is then gradually manoeuvred towards the lens to a final distance of a few tens of micrometers. A projection image of the nanotube through the lens aperture is shown in Fig. 11.2. In order to obtain a parallel beam, the lens is operated in the decelerating mode, applying a positive voltage at the lens entrance and keeping the lens exit as well as the sample at ground potential. The back side of the lens and of the lens holder have been shielded with a grounded conductive layer in order

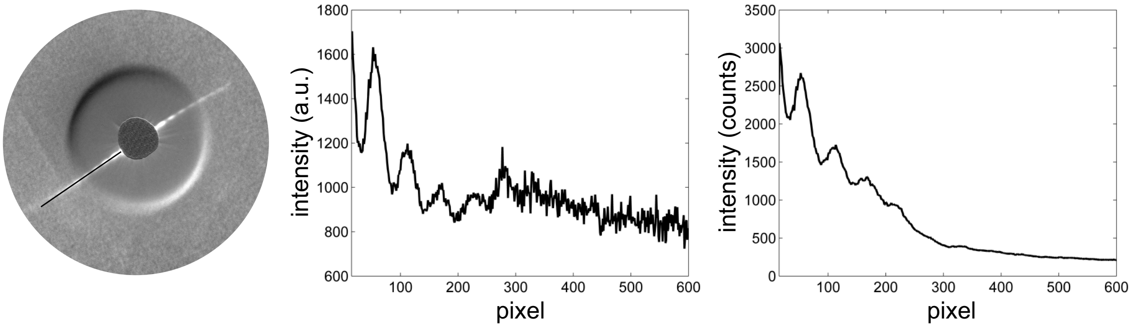


(a) (b)

Figure 11.3: *(a) Experimental diffraction pattern of a multi-walled carbon nanotube attached to a tip recorded with 57 eV electrons. (b) Profile along the equatorial line of the diffraction pattern showing intensity maxima up to the sixth order. The distance between the intensity maxima corresponds to a tube diameter of 9.4 nm.*

to avoid disturbing fields in the region between lens exit and sample-tip. During focusing of the beam the position of the sample is continuously adjusted to keep it well aligned to the beam. When a diffraction pattern becomes visible, emission tip and sample are simultaneously moved so that the diffraction pattern appears centred on the screen. A beam stop covers the zero order beam to avoid overexposure of the CCD. The lens-voltage and the emission current are finally adjusted to optimize the signal in the diffraction pattern. One of the diffraction patterns obtained in this way is shown in Fig. 11.3 (a). A sequence of 100 images has been summed in this case in order to reduce the noise level. The kinetic energy of the electrons at the sample amounted to $E_{\text{kin}} = 57\,\text{eV}$, corresponding to a wavelength of 1.6 Å. The lens voltage was set to U_L=54 eV. The total emission current during exposure was about 250 nA. Every image of the sequence was recorded with an acquisition time of 700 ms. The equatorial line in direction perpendicular to the tube direction with its intensity modulations is nicely visible in the displayed diffraction pattern. The intensity profile along the equatorial line shown in Fig. 11.3 (b) is symmetric, which means, that the nanotube does not alter the phase of the electron wave and thus constitutes a pure amplitude object. The distance between the intensity maxima along the equatorial line as it can be deduced from the profile corresponds to a tube thickness of 9.4 nm, which is a realistic value for a multi-walled carbon nanotube. Determination of the diameter of the very same tube by other means, e.g. TEM imaging, is not possible with this preparation method. From the distance of two maxima of the same order it can be inferred that the maxima of first and second order are hidden behind the beam stop. Thus, maxima up to the sixth order are clearly recognizable in the diffraction pattern. The resolution present in the diffraction pattern, i.e. the highest frequency where any diffraction signal is visible, is about 1 nm.

The bright circular *halo*-like region around the beam stop results from a detector artefact, always present around bright regions on the screen. Because of the long acquisition times of the CCD camera the halo appears much more pronounced in the diffraction patterns than in usual projection images and leads to a decreased contrast of the intensity modulations. By background correction, i.e. by dividing the diffraction pattern by a background image taken with the same parameters but with the sample moved out of the beam, the effect of the halo can be reduced as shown in Fig. 11.4. Unfortunately it can not be completely removed. Background correction additionally removes the features in the diffraction pattern originating from the lens aperture and thus improves the signal in the higher diffraction orders. Although the profile is corrupted by the edge that still remains due to the halo, there is at least one additional intensity maximum recognizable in the profile of the corrected image in Fig. 11.4 (e) compared to the same profile of the uncorrected image. Post-processing of the recorded images by background correction thus constitutes in general an easy and very useful method.



(a) (b) (c)

Figure 11.4: *For background correction the diffraction pattern is divided by a background image recorded with the same parameters but the sample moved out of the beam. The corrected image is displayed in (a). Higher order signal is enhanced by background correction, as it becomes obvious by comparison of the intensity profiles along the marked line of the corrected (b) and the uncorrected pattern (c). Since the noise in the background image has not been reduced by averaging, the noise level is higher in the corrected image.*

11.2.2 Experimental diffraction patterns of carbon nanotubes over holes in a support film

The preparation of carbon nanotubes over holes in a support film has several advantages compared to attaching a nanotube to a tip. First, the preparation procedure can be performed in a more routine way and examination of a larger number of nanotubes on one sample is possible. Besides, it is possible to post-accelerate the electrons towards the sample, as has been described in 10.3. Furthermore, the methods for preparing nanotubes on holey support films are extendible for the preparation of proteins attached to carbon nanotubes, as suggested in 10.5.5.

Samples with carbon nanotubes over holes in a support film have been prepared as described in 10.5.4. As mentioned and as it can also be inferred from the TEM image of a double-walled car-

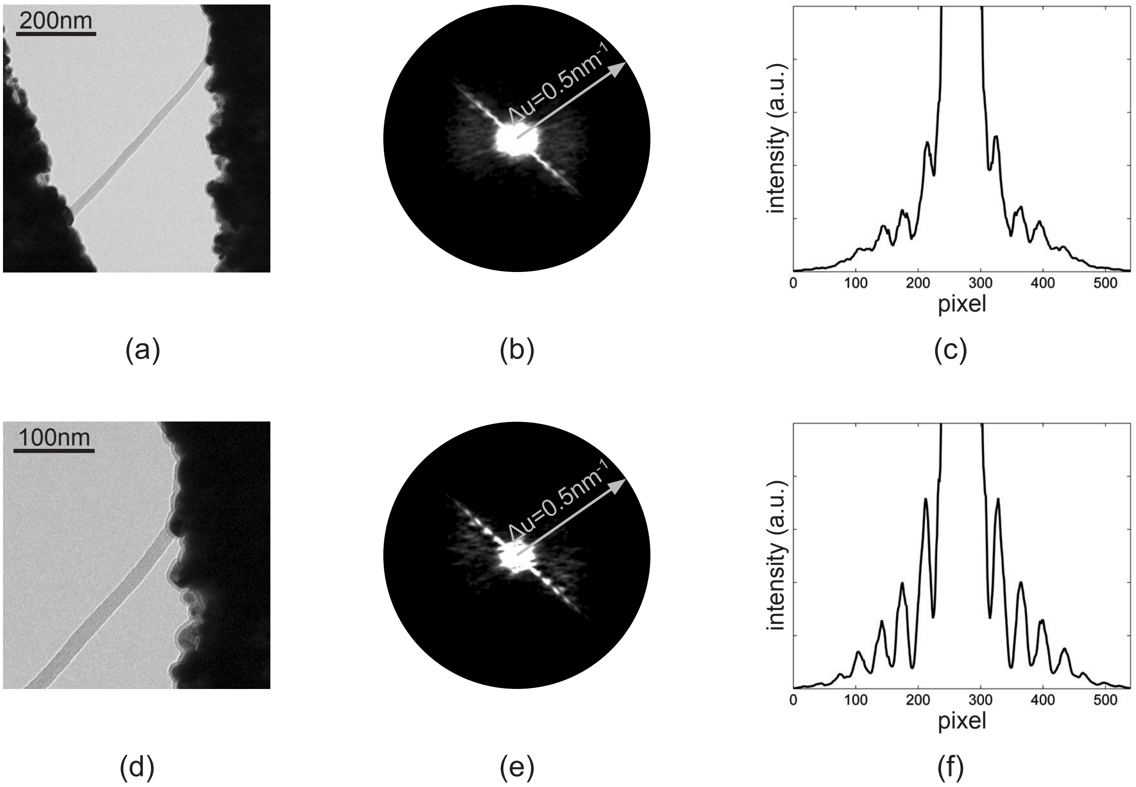
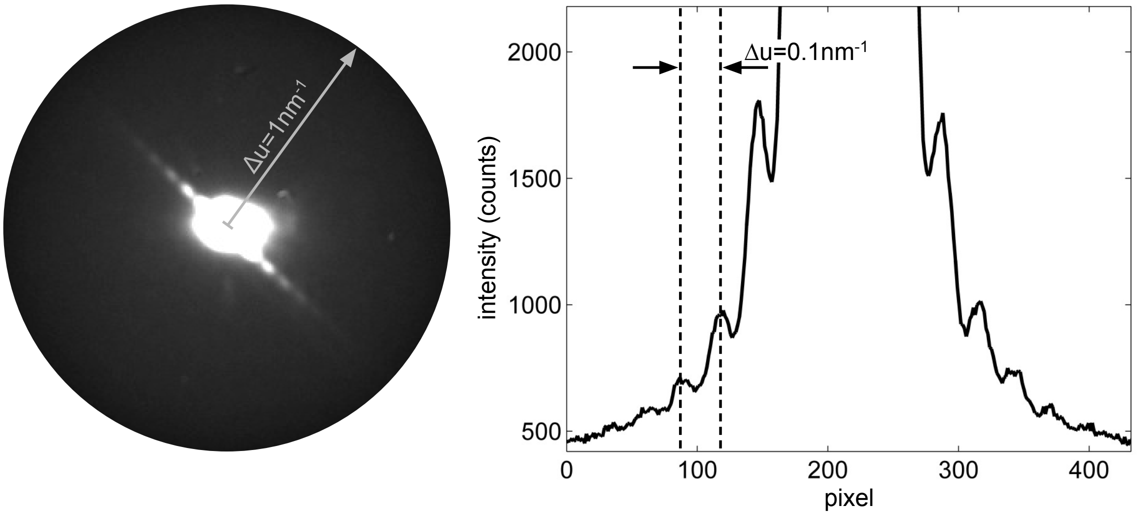


Figure 11.5: *Diffraction patterns calculated from TEM images. (a), (d) Different sections of a TEM image of a bundle of carbon nanotubes lying over a slit in a carbon support film coated with a 15 nm Ag layer. (b), (e) Calculated Fourier transforms of the TEM images smoothed with a Gaussian filter. (c), (f) Corresponding profiles along the modulations in the Fourier transforms.*

bon nanotubes sample in Fig. 11.5 (a), we mainly observe bundles of several tubes on the prepared samples while single nanotubes lying over holes are hardly found. The thickness of the nanotube bundles varies slightly along the 500 nm width of the slit and the bundles are not perfectly straight. The Fourier transform of such a structure exhibits intensity in direction perpendicular to the main tube direction. The distance between the modulation maxima corresponds to the mean tube diameter. Compared to the Fourier transform of a perfectly straight structure of constant thickness, the modulations appear less pronounced and the higher orders vanish. This is also illustrated in Fig. 11.5 (b) and (c), displaying the calculated squared modulus of the Fourier transform of the TEM image shown in Fig. 11.5 (a) and a profile along the equatorial line. To make it more comparable to the experimental diffraction patterns, the calculated Fourier transform has been blurred with a Gaussian filter. The modulation contrast along the equatorial line and the number of visible diffraction orders depends on which part of the nanotube is exposed to the beam. This is illustrated by Fig. 11.5 (e), where the Fourier transform of a section of the TEM image is displayed. The profile in Fig. 11.5 (f) shows that the modulation contrast is higher, since the part of the nanotube now contributing to the diffraction pattern is shorter and therefore rather resembling a straight structure of constant thickness. The information about the exact shape of the structure is contained in faint speckles hardly recognizable even in the calculated images.



(a) (b)

2

Figure 11.6: *(a) Experimental diffraction pattern as recorded of a CNT structure suspended over a slit in a support film showing modulations perpendicular to the longitudinal direction of the CNT bundle. (b) Profile along the modulations showing that the distance between the maxima corresponds to a thickness of the imaged structure of 10 nm.*

An experimental diffraction pattern of a nanotube structure is shown in Fig. 11.6. Positioning of the sample behind the lens aperture has been done in the projection mode of the microscope. The sample was then gradually moved towards the lens to a final distance of about 200 μm. During focusing of the beam, the sample position was readjusted in order to keep it aligned. The lens was operated in the decelerating mode, and beyond the lens the electrons were post-accelerated towards the sample. The displayed diffraction pattern has been recorded with electrons post-accelerated to a final kinetic energy of 186 eV, corresponding to a wavelength of 0.9 Å. The total emission current amounted to about 200 nA and the acquisition time of the CCD was set to 800 ms. The central region is significantly overexposed due to the non-scattered part of the beam. Perpendicular to the longitudinal direction of the structure modulations are visible; an intensity profile along this perpendicular direction is shown in Fig. 11.6 (b). The distance between the maxima corresponds to a tube diameter of about 10 nm. This is consistent with the TEM observations and confirms that in fact a bundle of nanotubes has been imaged. One can deduce from the intensity profile, that the intensity maxima of first and second order are hidden in the overexposed part of the image. Thus, modulations are visible up to the seventh order, which corresponds to what is expected from the calculated Fourier transforms. The outermost maximum corresponds to a resolution of 1.5 nm.

11.3 Conclusions

The diffraction patterns of carbon nanotube structures that could be recorded so far appear as expected from TEM images of the samples, taking into account the spatial resolution of the currently employed detector. Variations in thickness of the nanotube bundles being suspended over holes lead to fading of the intensity in the modulation maxima towards higher orders. With the presented preparation methods it appears unlikely to obtain individual carbon nanotubes stretched over holes such that they can be considered as structures being perfectly straight over a distance of several hundred nanometers. The principal advantage of coherent diffraction imaging is its applicability to objects of any shape, without any need for periodicity. The information about the object structure is contained in continuous speckles present in the entire diffraction pattern. In the experiments performed here, it is due to the limited spatial resolution and sensitivity of the detector that these speckles in the diffraction pattern can not be resolved, precluding subsequent phase retrieval. It is expected that the implementation of a detector with enhanced spatial resolution and sensitivity will allow to record diffraction patterns with a sampling rate sufficiently high for successful reconstruction.

Chapter 12

Discussion and outlook

With the presented experimental scheme and sample preparation methods it has been possible to obtain diffraction patterns of individual nanometer sized objects using a beam of coherent low-energy electrons. A key component of the designed experimental scheme consists in the micrometer sized electrostatic lens for collimating the divergent electron beam emitted from the source without destroying spatial coherence. Several methods have been developed, allowing routine fabrication of lenses with apertures of one micrometer in size. The down-scaled spherical aberrations of these microlenses, in combination with a source of atomic dimension emission area, permit to transform the initially divergent electron beam into a parallel one having a divergence angle of a few milliradians only. This allows coherent illumination of nanometer sized objects placed beyond the lens.

Sample preparation can be considered another challenging aspect of the experimental setup. Since, in view of subsequent image reconstruction, it is crucial to have the specimen surrounded by as little as possible disturbing support, it must be suspended over holes in a support film. A preparation method has been described here for placing carbon nanotubes over a few hundred nanometer sized holes in a carbon film. This method provided the first diffraction patterns of nanometer sized objects obtained with a coherent beam of low-energy electrons. Moreover, the described preparation procedure may be extended such that the carbon nanotubes serve as a template to which proteins can be attached by established biochemical methods. However, up to now, the applied preparation method resulted mostly in bundles of carbon nanotubes suspended over holes, while individual ones were rarely found. Therefore, developing alternative methods for the preparation of carbon nanotubes would constitute an improvement. Catalytically growing single-walled carbon nanotubes from nanoparticles, as for example described in [109], appears to be a method that could be adapted for our application since it yields clean samples of nanotubes of controllable diameter. Subsequent attachment of proteins has already proven to be possible [118].

With regard to first low-energy electron diffraction patterns of biological samples, DNA molecules constitute another suitable specimen, as they can be suspended over holes without the need for an additional template molecule. Preparation methods could be similar to the ones applied for LEEPS microscopy described in [9, 10]. When it is about finding a template for proteins, DNA could also be an alternative to carbon nanotubes, making use of the methods usually applied for fluorescent labelling of DNA.

A general question that might arise, is to what extent the structure of a biomolecule is preserved under ultra-high vacuum conditions. In cryo-EM, the samples are vitrified and kept frozen during imaging [119], ensuring that the structure of the molecule is not disrupted by crystalline ice. For the same reason, it has been suggested for the experiments on imaging of biological molecules at the XFELs to use spraying techniques in order to inject the proteins within a tamper layer of water or vitreous ice [120, 121]. However, it appears conceivable, that the water being part of the structure of a biological molecule remains bound to the molecule during freeze-drying the vitrified sample, hence preserving the molecular structure also under vacuum conditions.

The diffraction patterns of carbon nanotube structures recorded so far, show a resolution in the diffraction signal of about 1 nm. The sampling rate was not sufficiently high to allow subsequent reconstruction of the images by phase retrieval algorithms. Both is currently not limited by coherence issues but by the sensitivity of the detector and its spatial resolution. It is expected, that once the newly constructed detector is implemented having considerably improved spatial resolution and enhanced sensitivity, the recording of oversampled diffraction patterns containing information about the specimen at the sub-nanometer resolution will become possible. Images of the samples shall then be obtained by applying phase retrieval algorithms as they have also been sketched here. Their performance when applied to experimental low-energy electron diffraction patterns is yet to be seen, concerning especially the noise level in the images and the handling of the missing data about the low-resolution part of the diffraction pattern. The best way how the constraints have to be imposed in order to reconstruct the complex valued images, still remains to be figured out.

As has been pointed out, low-energy electrons are the only known radiation with wavelengths in the Ångstrom regime that does not destroy an individual biological molecule before sufficient elastic scattering events have been recorded. This allows imaging their structure without the need to average over a large number of copies of the specimen. Hence, the successful recording of first coherent low-energy electron diffraction patterns of nanometer sized objects constitutes a major step towards imaging biomolecules. Revealing the structure of individual biomolecules, coherent diffraction imaging with low-energy electrons could contribute to solve the puzzles of structural biology.

Appendix A

Publication

Ultramicroscopy 110 (2010) 1148–1153

Contents lists available at ScienceDirect

Ultramicroscopy

journal homepage: www.elsevier.com/locate/ultramic

Fabrication and characterization of low aberration micrometer-sized electron lenses

E. Steinwand*, J.-N. Longchamp, H.-W. Fink

Institute of Physics, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland

A R T I C L E I N F O

Article history:
Received 11 February 2010
Received in revised form
1 April 2010
Accepted 20 April 2010

Keywords:
Microlens
Spherical aberrations
Electron optics
Low-energy electrons
Field emission

A B S T R A C T

Intrinsic spherical aberrations of electron lenses have been the major resolution limiting factor in electron microscopes for several decades. While effective correctors have recently been implemented, an alternative to correct these aberrations is to circumvent them by scaling down lens dimensions by several orders of magnitude. We have fabricated electrostatic lenses exhibiting one micrometer diameter apertures and evaluated their beam forming properties against predictions from numerical ray tracing simulations. It turns out that it is routinely possible to shape a paraxial low-energy electron beam by such micron-sized lenses. Beam profiles have been measured both at a distant detector as well as in a plane close to the lens. It is shown that the lens can form a parallel beam extending no more than 800 nm from the optical axes at a distance of 200 μm beyond the lens exit. We believe that these findings constitute a prerequisite to derive novel tools for high resolution microscopy using low-energy electrons.

1. Introduction

Back in 1936 already, Scherzer [1] recognized that rotational symmetric electrostatic electron lenses suffer from intrinsic aberrations which have limited the resolution of conventional electron microscopes for more than 50 years. These intrinsic limitations can only be overcome by introducing elaborate electron optical elements, like multi-pole correctors and mirrors, into the path of an electron beam. Actually building and implementing such correctors has only recently been achieved [2,3] and revolutionized modern electron microscopy design. An alternative approach towards minimal aberrations relies on the concept of scaling down both, electron source and lens dimensions. In scaling down the size of an electrostatic lens, while keeping the electrode potentials unchanged, the shape of equipotentials and electron trajectories is maintained; they both just undergo the same similarity transformation as the lens geometry. As a consequence, the spherical aberration coefficient, is directly proportional to the lens size. While several attempts have already been undertaken in this direction [4–6], routine operation of low aberration micron-scale lenses and their application in scientific instruments are still lacking. While we routinely employ electron sources exhibiting an ultimate emission area of atomic dimension [7] it is now a matter of scaling down a lens by about four orders of magnitude and positioning it with nanometer precision in front of the source.

2. Microlens fabrication

2.1. Basic design criteria

The simplest type of electrostatic lens consists of two parallel planar electrodes with two concentric apertures of not necessarily the same size. The electrodes must be separated by an insulating material to maintain a voltage between them, resulting in a focusing electrostatic field distribution at the apertures as illustrated in Fig. 1. A focusing effect is generated independent of the polarity of the applied lens-voltage. A lens size as small as possible appears preferable in order to minimize spherical aberrations. Considering established microfabrication techniques and available materials, lens dimensions in the range of one micrometer appear sensible. Micrometer lens dimensions shall also ensure easy positioning of the lens in an electron–optical system using conventional nanopositioning devices based on piezo-electric manipulators.

2.2. Materials and methods

We have developed several lens fabrication methods comprising various microfabrication steps, evaporation methods and materials for the insulating layers and lens electrodes. The result

---

* Corresponding author.
E-mail address: elvira@physik.uzh.ch (E. Steinwand).

0304-3991/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.ultramic.2010.04.013

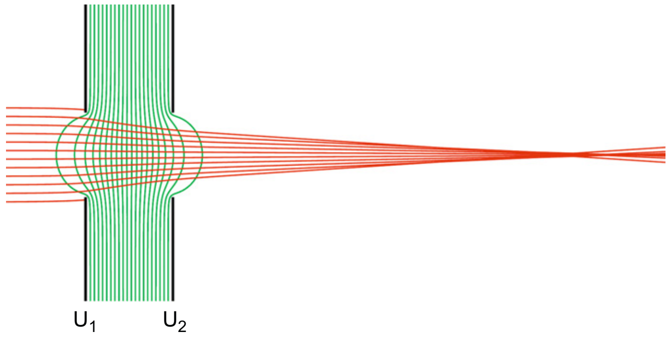


Fig. 1. Operating principle of a two electrode aperture lens for an accelerating electrical field between the electrodes. Equipotential lines and electron trajectories have been calculated for $U_1 = 100$ V and $U_2 = 350$ V for an initial electron energy of 100 eV.

of all methods is a lens structure similar to the one depicted in Fig. 2(d) featuring a 1 μm thick insulating layer, two electrodes of several ten nanometers thickness with one micrometer diameter apertures in them. The electrodes must be freestanding around the apertures to avoid charging effects of the insulating walls while the electron beam impinges on the lens. The fabrication of one particular type of microlens, showing good performance in terms of stability and cleanliness, will be described here in some detail. As a starting material for the fabrication we use commercially available silicon nitride (SiN) membranes of one micrometer thickness. They serve as insulating layers between the two lens electrodes and exhibit a measured breakdown voltage around 320 V, high enough to focus a beam of electrons with kinetic energies in the range of 100 eV. The initial fabrication step consists in electron beam evaporation of a roughly 30 nm thick layer of amorphous carbon on either side of the SiN membrane (Fig. 2(a)). Next, a focused gallium ion beam (FIB) [8] is used to remove the top carbon as well as the SiN layer within a circular region of 5 μm in diameter, whereas the carbon layer at the bottom side of the structure remains unchanged (Fig. 2(b)). For the following process step carbon is first evaporated onto a mica sheet. Subsequently, the carbon film is floated off onto a clean water surface [9] and can thus be deposited onto the upper side of the structure (Fig. 2(c)). As the final step, a hole of 1 μm diameter is ion-milled through the freestanding parts of both carbon films (Fig. 2(d)). A SEM image of such a final lens structure is shown in Fig. 2(e).

3. Experimental implementation and qualitative tests of the microlens

3.1. Experimental setup

Lenses fabricated as described above have been tested in an ultra-high vacuum system designed for experiments with coherent low-energy electrons. A W(1 1 1) field emission tip is used as a source for a divergent electron beam of high spatial and temporal coherence. Typical emitter currents are in the 10–200 nA range and the kinetic energy of the electrons at the lens entrance is well below 200 eV. The electron detector consists of a microchannel plate (MCP) followed by an electroluminescent layer on the vacuum side of a fiber optic plate (FOP). At the ambient pressure side of the FOP a CCD camera collects the emitted light. The detector resolution has been measured to be around 120 μm. A dedicated holder allowing for rapid vacuum transfer of microlenses fixes the position of the lens (Fig. 3). The distance between lens and detector amounts to 75 mm. The electron source is mounted onto an x–y–z piezo-stage for precise alignment with the lens aperture.

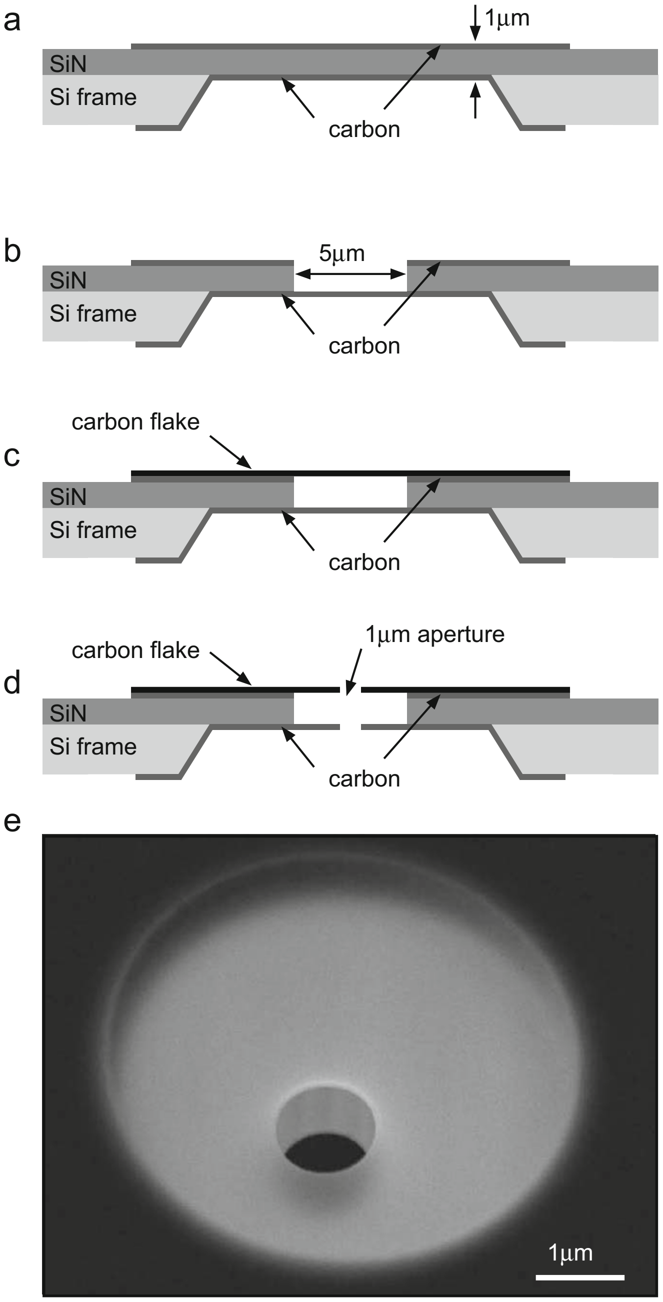


Fig. 2. Schematic drawings of the various steps for lens fabrication and an SEM image of the final lens structure. (a) Carbon deposition on either side of a SiN membrane. (b) Removing the upper carbon and the SiN layer within a circular 5 μm diameter region using a FIB. (c) Covering of the structure with a carbon flake. (d) Milling a 1 μm diameter aperture through the freestanding carbon layers using the FIB. (e) SEM image of the fabricated microlens, recorded at a tilt angle of 30°. The penetration of the 12 keV electrons used in the SEM, allows recognizing the 5 μm diameter circular region where the carbon electrodes are freestanding. In this case, the lens aperture of 1 μm in diameter, is not concentric with the 5 μm region but the distance to the SiN walls is large enough to prevent charging effects when implemented as a lens.

For the experiments described here, the source to lens distance varied between 5 and 30 μm, leading to kinetic energies of the electrons at the lens entrance between 60 and 150 eV.

3.2. Lens operation and overall performance

If the lens electrodes are both at ground potential and a negative voltage is applied to the emitter tip, a projection image of the second lens aperture is visible at the screen. Its magnification can be varied by changing the source–lens distance. Examples of such electron projection images are shown in Fig. 4(a) and (e). Once a voltage is applied between the two lens electrodes, a focusing effect of the lens is observed. To ensure a field-free region beyond the lens, the second electrode is always kept at ground potential. The voltage applied at the first lens electrode is altered together with the voltage at the emitter tip to maintain the kinetic energy of the electrons at the lens entrance and the emission current constant. The polarity of the voltage at the first lens electrode can either be such that the electrons are decelerated when passing the lens (positive voltage), or such that they are accelerated (negative voltage). The effect of the lens is illustrated in Fig. 4 for both modes of operation. While increasing the voltage between the two lens electrodes, one first observes that the projection image of the lens aperture at the screen decreases (Fig. 4(b) and (f)). A further increase of the lens voltage reduces the divergence angle more and more until the electrons form a minimal spot at the detector (Fig. 4(c) and (g)). With still higher lens voltages, a crossover located between lens and detector is achieved (Fig. 4(d) and (h)). The distance between lens and crossover can even get smaller than the source–lens distance, resulting in an enlarged divergence angle of the beam as illustrated in Fig. 4(d).

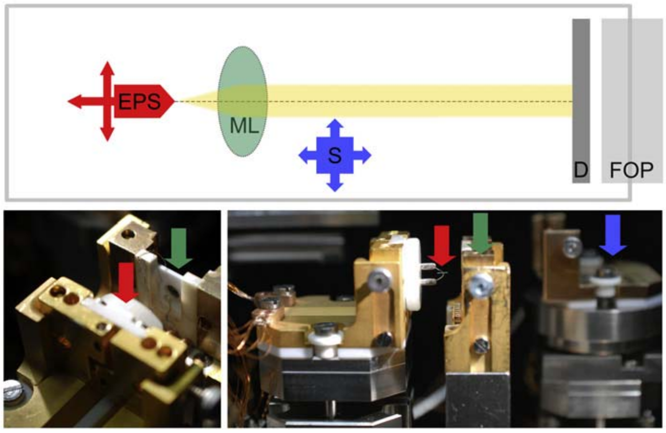


Fig. 3. Top: schematic of the experimental UHV setup. The electron point source (EPS) is mounted onto an *x*–*y*–*z* piezo-stage for nanometer precision alignment with the center of the one micrometer diameter aperture of the microlens (ML). Beyond the microlens, a second *x*–*y*–*z* piezo-stage carries a sample holder (S) to be moved into the beam. The beam profile is measured at a detector (D) placed at 75 mm beyond the lens. A fiber optic plate (FOP) transfers the image from the UHV to the ambient pressure side where it is captured by a CCD camera. Bottom: two views into the experimental chamber showing the electron point source (red arrow), the microlens holder (green arrow) and the sample holder (blue arrow).

4. Quantitative lens characterization

To evaluate whether downscaling the spherical aberrations by scaling down the size of the lens could actually be realized, the magnitude of spherical aberrations must be determined. Two quantities which are directly related to spherical aberrations have been measured. In a second step they were compared to the values obtained from ray tracing simulations, assuming an idealized lens deteriorated only by intrinsic spherical aberrations. The order of magnitude of the spherical aberration coefficients de facto realized in the experiments could thus be deduced.

4.1. Experimental methods

4.1.1. Beam profile at a distant detector

An experimentally easily accessible quantity is the image size of the electron source in the plane of the electron detector which is the smallest achievable spot size at the detector. As the distance from the lens to the detector plane is large compared to the diameter of the lens aperture and to the source–lens distance, a minimal spot size at the detector is achieved when the beam leaves the lens almost parallel. Instead of determining spherical aberrations from the size of the spot obtained when a parallel

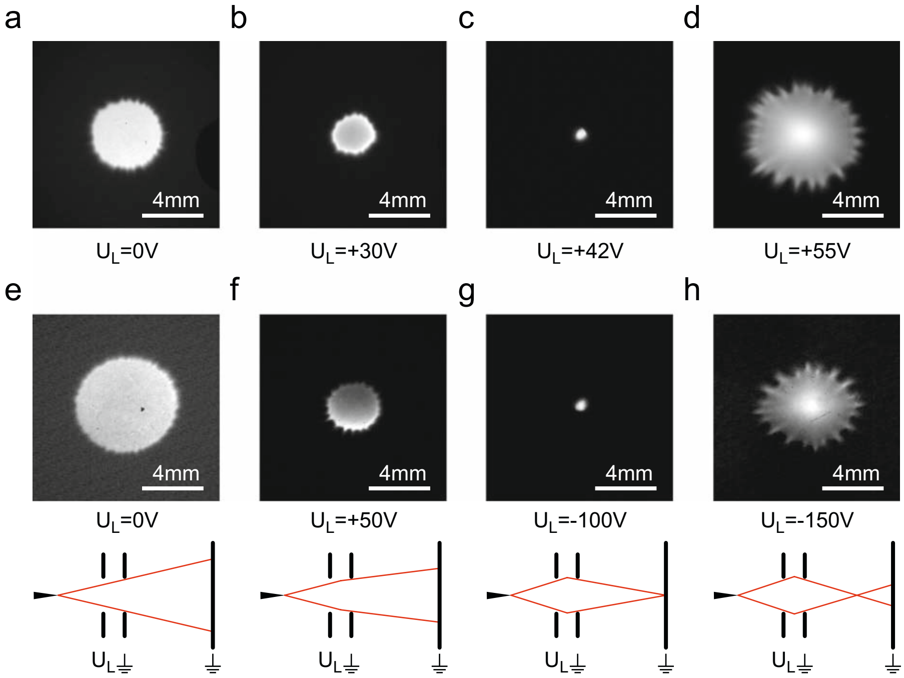


Fig. 4. Focusing series for the decelerating (top) and accelerating (middle) mode of the lens with corresponding focusing situations (bottom). The distance between the virtual electron source and the first lens electrode amounts to 16 μm for the decelerating and 11 μm for the accelerating mode with kinetic energies of 93, respectively, 95 eV at the lens entrance.

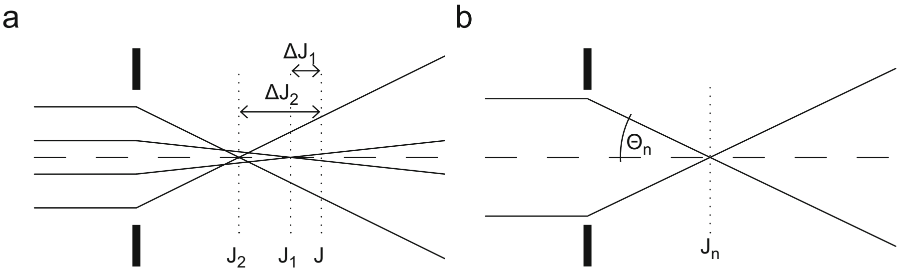


Fig. 5. (a) Schematic drawing showing two rays with different initial distances from the optical axis which are focused at different positions J_1 and J_2. The paraxial focus is denoted as J. ΔJ_1 and ΔJ_2 are called the longitudinal spherical aberration of the two rays. (b) The position of the focus J_n depends on the semi-angular aperture Θ_n of the ray.

incoming beam is focused, we did it the other way round. We instead measured the minimal achievable divergence angle when collimating a beam emitted by an almost perfect point source. Thus, the minimal spot size at the detector is directly related to the spherical aberrations of the lens.

The minimal achievable spot size was measured for the decelerating mode of the lens. Each measurement started with taking a projection image of the lens aperture. This allows determining the distance between the virtual electron source and the first lens aperture [10]. Next, the lens voltage was adjusted for obtaining a minimal spot at the distant detector. In order to get the size of the minimal spot and the projection image, the contour line corresponding to half the maximum intensity value in the image was determined. An ellipse was then fitted to the contour line and the spot diameter was taken as the mean of its minor and major axis. It turns out, that the average spot size at the detector varies from 0.5 to 1.1 mm measured with a detector resolution of 120 μm. Thus, assuming a Gaussian spot profile the broadening of the spot imposed by the finite detector resolution amounts to $<3\%$ and will not be accounted for in the following. In addition to spherical aberrations there are several other intrinsic effects contributing to the size of the spot visible at the detector, namely chromatic aberrations, diffraction of the beam at the lens aperture and finite source size. For an estimation of chromatic aberrations the energy spread of 0.1% of the electrons [12] must be taken into account as well as the stability of the lens voltage. Ray tracing simulations show, that this leads to an enlargement of the spot size at the detector by about 15 μm. We estimated the influence of diffraction for the decelerating mode of the lens by assuming a parallel beam being diffracted at the second aperture of the lens. This leads to a width of the central maximum of the diffraction pattern of 15 μm for 50 eV electrons. The magnified image of the virtual electron source contributes with $<10\,\mu$m to the spot size, as the size of the virtual electron source is well below 1 nm; in fact it has been measured to be of atomic dimension [7]. We can thus conclude that chromatic aberrations, diffraction at the lens aperture and finite source size are negligible contributions to the minimal spot size at the detector.

4.1.2. *Beam profile at* 200 μm *distance beyond the lens exit*

A more direct way to determine aberrations is to measure the beam diameter in a plane much closer to the lens than the 75 mm distant detector plane. This has been done by scanning a sharp edge perpendicular to the optical axis through the beam. The beam diameter in the plane of the edge is given by the displacement of the edge from the position where all electrons are blocked to the position where all electrons pass the edge. We fabricated such sharp edge by ion-milling a rectangular window into a 20 nm thick carbon foil, opaque for low-energy electrons. The sample was mounted onto a movable piezo-stage (Fig. 3) and the window was positioned into the electron beam 200 μm beyond the lens. Precise motion of the edge perpendicular to the optical axis is realized with a piezo-scanner exhibiting a scan-range of 4 μm. While moving the sample with a scan speed of typically 800 nm/s, the total intensity at the detector was recorded. The lens voltage was then adjusted for the steepest slope in the intensity versus edge position profile while repeatedly scanning through the beam.

4.2. *Ray tracing simulations*

Calculation of electrostatic fields and electron trajectories have been done using the software package SIMION 8.0[1] to solve the Laplace equation using the finite-difference method. The lens was modelled as two planar electrodes separated by 1 μm and exhibiting concentric apertures of 1 μm in diameter with field-free regions on either side of the lens. Cylindrical symmetry reduces the computational effort since the Laplace equation must only be solved in two dimensions. The electrical field distributions were calculated on a $15\,000\times1000$ array with a grid size of 5 nm. Ray tracing was done with a fourth order Runge–Kutta algorithm implemented in SIMION.

4.2.1. *Calculation of spherical aberration coefficients*

Spherical aberrations cause paraxial and marginal rays to be focused at different positions, as illustrated in Fig. 5. The distance between the paraxial focus J and the focus J_n of a ray passing the lens at a larger distance from the optical axis is called longitudinal spherical aberration ΔJ_n of the ray, see Fig. 5(a) for denotation. In the case of a parallel incoming beam, the longitudinal spherical aberration of a ray ΔJ_n can be expressed as a power series of the ray's semi-angular aperture Θ_n (see also Fig. 5(b)). As ΔJ_n is a symmetric function of Θ_n, the odd power terms vanish, which leads to

$$\Delta J_n = C_s\Theta_n^2 + c_4\Theta_n^4 + c_6\Theta_n^6 + \cdots,$$

with the coefficient C_s being the primary spherical aberration coefficient [13]. The latter was calculated by combining ray tracing simulations with a least square fit, similar to methods described elsewhere [14]. Calculations were carried out for a parallel incident ray, both for the accelerating as well as for the decelerating mode of the lens. For this, a parallel beam of 110 electrons has been generated, such that the lens aperture was completely filled. For each particle, the focus position J_n and the elevation angle Θ_n were computed. The position of the paraxial focus J was approximated by the focus of the ray with an initial distance to the optical axis of 5nm. In this way, the longitudinal aberration $\Delta J_n = J - J_n$ was determined for each particle. A polynomial of sixth order was then fitted to the values $\Delta J_n(\Theta_n)$, yielding the primary aberration coefficient C_s. The described

---

[1] SIMION 8.0, Scientific Instrument Services, Inc., 2003–2006.

procedure was applied to a series of lens voltages for the decelerating and accelerating mode, respectively. In the following, voltages are always taken relative to the potential where the kinetic energy of the electrons is zero. Thus, the lens is completely described by its geometry and the voltage ratio between the two electrodes [13]. The dependence of C_s on the voltage ratio for the accelerating and decelerating mode are shown in Fig. 6 for voltage ratios related to realistic experimentally accessible values. The aberration coefficients C_s are in the range of millimeters for both modes and decrease with increasing focusing strength of the lens. Here, C_s denotes the aberration coefficient referred to the image side. Thus, the value for the lateral size of the focal spot in the plane of least confusion d_s can be calculated using the relation $d_s = 0.5 C_s \Theta^3$, where Θ is the angular aperture at the image side of the lens [11]. For a crossover several ten micrometers away from the lens, this yields a spot size of several 10 nm.

4.2.2. *Calculation of the spot size at a given distance*

For comparison with experimental values the size of the image of the electron source was calculated in the detector plane 75 mm beyond the lens and in the plane located 200 μm beyond the lens. As the main contribution to the finite spot size in a plane distant to the lens is due to spherical aberrations, other factors influencing the spot size, as there are chromatic aberrations, finite source size and diffraction effects, were neglected in the simulations. A monochromatic point source was thus assumed. Since typical experimental tip to lens distances are larger than 5 μm, the electrostatic field near the first lens-aperture and the strong field close to the field emission tip do not influence each other, so that they can be treated as two separated electron–optical components. The electron trajectories in the immediate vicinity of the field emission tip are slightly curved as most of the potential drops in front of the tip. However, further away from the tip, at the position of the lens, the trajectories are straight lines. At the lens entrance they seem to emanate from a virtual source located a short distance behind the physical tip apex [10]. There is thus no need to include the field emission tip in the electrode array. Instead, the region at the entrance side of the lens can be assumed as field-free. Electrons originate at a point corresponding to the position of the virtual source and propagate straight with given kinetic energies and divergence angles towards the lens. As the region behind the lens is field-free, the beam diameter in the planes of interest up to the 75 mm distant detector was calculated by extrapolating the trajectories. In accordance with the experimental conditions, a sufficiently large divergence angle was selected to ensure that the electrons completely fill the lens aperture. The distance between subsequent electrons in the plane of the lens aperture was set as close as possible, that is to one grid unit corresponding to 5 nm. For several distances between electron source and first lens-aperture the spot size in the plane under consideration was calculated for a set of various lens voltages. In this way, the minimal possible spot size was computed as a function of source–lens distance for both planes evaluated in the experiments, which is at 75 mm and 200 μm beyond the microlens.

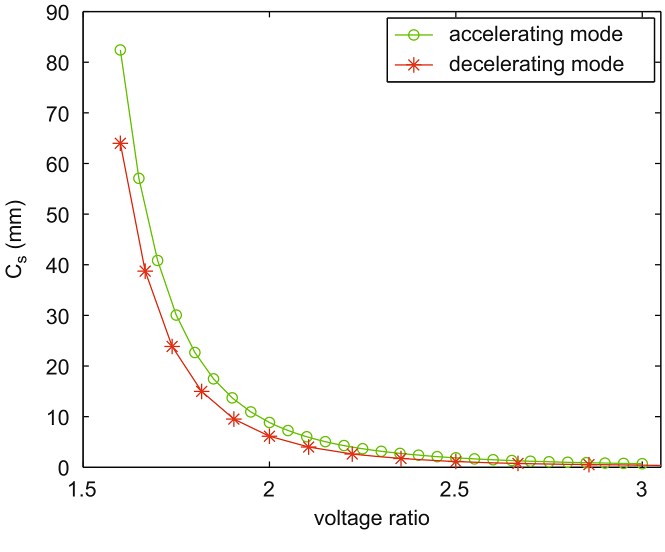


Fig. 6. Primary aberration coefficient C_s as a function of the voltage ratio for the decelerating and accelerating mode of the lens.

5. Results and discussion

5.1. *Minimal spot size at the 75 mm distant detector*

Measured spot sizes at the detector as a function of source–lens distance are depicted in Fig. 7 together with the corresponding simulated values. Simulations predict a minimal spot size at the detector between 0.2 and 1.1 mm in diameter at the detector for source–lens distances ranging from 5 to 30 μm. As expected, the increase of the divergence angle with decreasing source–lens distance leads to a larger spot size. The simulated values exhibit some fluctuations which are due to the limited precision in the determination of the lens-voltage required to form a minimal spot. Two series of experimental measurements of the spot size at the detector are also plotted in Fig. 7. For series 1 several measurements have been carried out with one and the same microlens, fabricated in the manner described above. In contrast to this, the eight data points of series 2 have been obtained from measurements with eight different microlenses. Some of those lenses have been produced by methods different from the one described here; however, all of them exhibit the same geometry. The measured values for the minimal spot size at the detector show the predicted dependency. The experimental data points, however, are shifted by 0.3 mm towards higher values compared to the simulated ones. The measured spot sizes assume values between 0.5 and 1.1 mm corresponding to a divergence angle of the collimated beam between just 6 and 14 mrad.

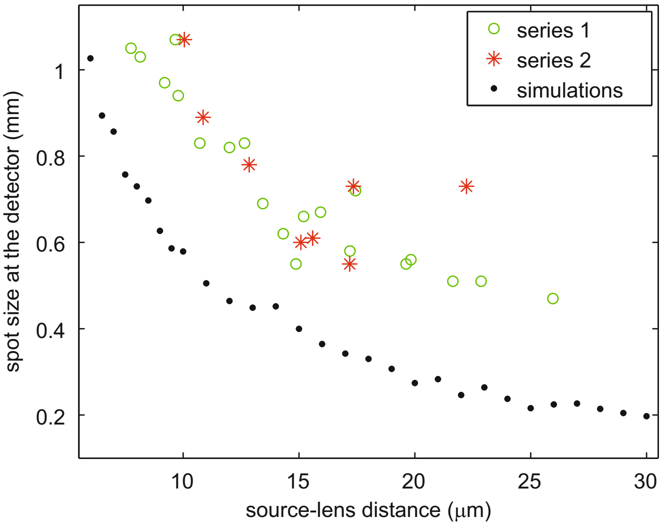


Fig. 7. Comparison between simulated and experimental values for the minimal spot size at the detector. Data points of series 1 have been obtained from measurements with one individual lens while the data for series 2 correspond to measurements with eight different lenses.

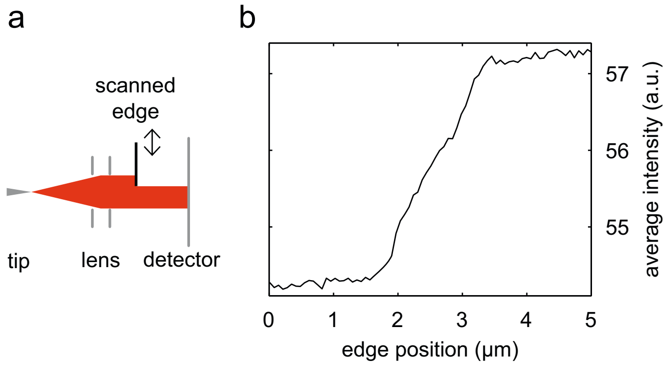


Fig. 8. (a) To measure the beam diameter 200 μm beyond the lens, a micromachined edge is scanned through the beam while the total intensity at the screen is recorded. (b) Measured intensity at the screen as a function of edge position. The source–lens distance amounts to 14 μm. A beam diameter of 1.8 μm is deduced from the width of the step profile.

The experimental values of series 2 obtained with different lenses exhibit similar characteristics. This implies that microlenses with comparable quality can now routinely be fabricated in a reproducible way. Moreover, as different lens fabrication methods result in similar lens performances, they appear to be equivalent with respect to the focusing properties of the lenses.

5.2. *Minimal spot size in a plane* 200 μm *beyond the lens*

The simulated spot-sizes in the plane 200 μm beyond the lens vary from 3 μm diameter for a source–lens distance of 6 μm down to 0.6 μm diameter for a source–lens distance of 30 μm. Experimental values have been obtained as described above (see also Fig. 8(a)). For a source–lens distance of 14 μm and 92 eV electrons at the lens entrance, a lens-voltage of 48 V was found to generate a minimal spot. In Fig. 8(b) the total intensity at the screen, averaged over all pixels of the CCD chip, is plotted versus the position of the edge moved perpendicular to the beam with a scan speed of 800 nm/s. From the intensity profile a beam diameter of 1.8 μm has been derived. Corresponding simulations assuming identical conditions reveal a value of 1.2 μm. As the distance of 200 μm is still large compared to the 1 μm diameter of the lens apertures, focusing of the beam towards this plane is impossible. Instead, the beam diameter is minimal when the beam leaves the lens as parallel as possible. Thus, the focusing conditions for a minimal beam diameter are just as in the case of a 75 mm distant detector as described and discussed above. In fact, the values for the beam divergence angle obtained by the two methods are in very good agreement; they amount to 9 mrad for a source–lens distance of 14 μm.

6. Conclusions and outlook

We have shown that electrostatic lenses of micrometer dimensions can readily be fabricated and exhibit reproducible performance. Comparisons with ray tracing simulations show that the concept of scaling down spherical aberrations by decreasing lens dimensions can in fact be realized. Effects like misalignment of tip and lens, residual vibrations and ac-magnetic fields, deviations from perfect lens-symmetry and contaminations of the apertures have not been considered at all in our simulations. Therefore, the experimental values for the spot size are somewhat larger than the values predicted by simulations. Although these effects do reduce lens performance slightly, their magnitude turns out to be small compared to the achieved reduction of spherical aberrations by down-scaling lens dimensions. As a result, a simple micrometer-sized lens exhibits aberration coefficients comparable to those of high performance objective lenses found in modern electron microscopes [15]. While the bare number of the aberration coefficient is just one aspect, an even more important feature of the scaling concept is the fact that the electron beam is always kept close to the optical axes. The coherent divergent beam originating from the electron point source is modified early on by the lens and thus never spreads out to macroscopic dimensions. The nearly parallel beam beyond the microlens deviates <1 μm from the optical axes at a distance of 200 μm beyond the lens. This implies that all following electron optical devices, as an objective lens for example, perceive a micron dimension paraxial electron beam with a broadening of just 10 mrad. No beam limiting aperture to improve resolution but decrease brightness at the same time would consequently be needed for an objective lens positioned beyond the microlens. Furthermore, since the microlens is combined with a coherent electron point source here, imaging technologies relying on the wave character of electrons, like holography or coherent diffraction, appear feasible. The atomic dimension emission area of the source in combination with the microlens appears to also be favorable when it comes to creating a focus by means of an objective lens. Since the focus is nothing but the image of the primary source, there would be no need to obtain a demagnified image of the primary source. A sharp focussed beam could thus be obtained and directed towards a sample placed at a convenient large working distance. In having shown the performance of such microlenses, it is now a matter of exploring exciting applications in imaging with coherent low-energy electrons.

Acknowledgment

We would like to thank the Swiss National Science Foundation for financial support.

References

[1] O. Scherzer, Z. Phys. 101 (1936) 593–603.
[2] O.L. Krivanek, N. Dellby, A.R. Lupini, Ultramicroscopy 78 (1999) 1–11.
[3] H.H. Rose, J. Electron. Microsc. 58 (2009) 77–85.
[4] T.H.P. Chang, D.P. Kern, L.P. Muray, J. Vac. Sci. Technol. B 8 (1990) 1698–1705.
[5] G.M. Shedd, H. Schmid, H.-W. Fink, Ultramicroscopy 48 (1993) 43–48.
[6] G.M. Shedd, H. Schmid, P. Unger, H.-W. Fink, A.D. Dubner, Rev. Sci. Instrum. 64 (1993) 2579–2584.
[7] A. Degiovanni, R. Morin, in: Electron Microscopy, Proceedings of the ICEM 13, vol. 1, Paris, 1994, pp. 331–332.
[8] J. Orloff, M. Utlaut, L. Swanson, High Resolution Focused Ion Beams, Kluwer Academic/Plenum Publisher, New York, 2003.
[9] M.J. Dykstra, L.E. Reuss, Biological Electron Microscopy; Theory, Techniques and Troubleshooting, Springer, Netherlands, 2003.
[10] P.W. Hawkes, E. Kasper, Principles of Electron Optics, vol. 2, Academic Press, London, 1996.
[11] P.W. Hawkes, E. Kasper, Principles of Electron Optics, vol. 1, Academic Press, London, 1996.
[12] R. Gomer, Field Emission and Field Ionization, Harvard University Press, Cambridge, Massachusetts, 1961.
[13] O. Klemperer, Electron Optics, Cambridge University Press, Cambridge, 1953.
[14] M. Oral, B. Lencová, Ultramicroscopy 109 (2009) 1365–1373.
[15] L. Reimer, H. Kohl, Transmission Electron Microscopy, Springer, New York, 2008.

Bibliography

[1] R. Feynman, *There's Plenty of Room at the Bottom*, Engineering and Science **13** (1960).

[2] K. Wüthrich, *Protein Structure Determination in Solution by NMR Spectroscopy*, J. Biol. Chem. **265**, 22059–22062 (1990).

[3] J. C. H. Spence and R. B. Doak, *Single Molecule Diffraction*, Phys. Rev. Lett. **92**, 198102 (2004).

[4] M. van Heel et al., *Single-particle electron cryo-microscopy: towards atomic resolution*, Q. Rev. Biophys. **33**, 307–369 (2000).

[5] E. Nogales, S. G. Wolf, and K. H. Downing, *Structure of the $\alpha\beta$ tubulin dimer by electron crystallography*, Nature **291**, 199–203 (1998).

[6] B. Böttcher, S. A. Wynne, and R. A. Crowther, *Determination of the fold of the core protein of hepatitis B virus by electron cryomicroscopy*, Nature **386**, 88–91 (1997).

[7] R. Neutze, R. Wouts, D. van der Spoel, E. Weckert, and J. Hajdu, *Potential for biomolecular imaging with femtosecond X-ray pulses*, Nature **406**, 252–257 (2000).

[8] H. N. Chapman, *X-ray imaging beyond the limits*, Nat. Mater. **8**, 299–301 (2009).

[9] M. Germann, T. Latychevskaia, C. Escher, and H.-W. Fink, *Nondestructive Imaging of Individual Biomolecules*, Phys. Rev. Lett. **104**, 095501 (2010).

[10] H.-W. Fink, H. Schmid, E. Ermantraut, and T. Schulz, *Electron holography of individual DNA molecules*, J. Opt. Soc. Am. A **14**, 2168 (1997).

[11] D. Gabor, *A new microscopic principle*, Nature **161**, 777–778 (1948).

[12] O. Scherzer, *Über einige Fehler von Elektronenlinsen*, Z. Phys. **101**, 593–603 (1936).

[13] O. Scherzer, *Sphaerische und chromatische Korrektur von Elektronenlinsen*, Optik **2**, 114–132 (1947).

[14] O. L. Krivanek, N. Dellby, and A. R. Lupini, *Towards sub-Å electron beams*, Ultramicroscopy **78**, 1–11 (1999).

[15] O. L. Krivanek, P. D. Nellist, N. Dellby, M. F. Murfitt, and Z. Szilagyi, *Towards sub-0.5Å electron beams*, Ultramicroscopy **96**, 229–237 (2003).

[16] H. H. Rose, *Historical aspects of aberration correction*, J. Electron. Microsc. **58**, 77–85 (2009).

[17] T. H. P. Chang, D. P. Kern, and L. P. Muray, *Microminiaturization of electron optical systems*, J. Vac. Sci. Technol. B **8**, 1698–1705 (1990).

[18] T. H. P. Chang, D. P. Kern, M. A. McCord, and L. P. Muray, *A scanning tunneling microscope controlled field emission microprobe system*, J. Vac. Sci. Technol. B **9**, 438–443 (1991).

[19] G. M. Shedd, H. Schmid, and H.-W. Fink, *Building a micrometer-scale electrostatic lens by hand*, Ultramicroscopy **48**, 43–48 (1993).

[20] G. M. Shedd, H. Schmid, P. Unger, H.-W. Fink, and A. D. Dubner, *Combinations of point source electron beams and simple electrostatic lenses: Initial demonstrations of micron-scale lenses*, Rev. Sci. Instrum. **64**, 2579–2584 (1993).

[21] H. Schmid, H.-W. Fink, C. Schiller, and T. L. van Rooy, *Focusing properties of micron-sized immersion lenses*, Rev. Sci. Instrum. **67**, 375–377 (1996).

[22] E. Steinwand, J.-N. Longchamp, and H.-W. Fink, *Fabrication and characterization of low aberration micrometer-sized electron lenses*, Ultramicroscopy **110**, 1148–1153 (2010).

[23] L. Zambov, G. Peev, V. Shanov, and S. Drumeva, *Composition, structure and properties of silicon nitride films grown from dichlorosilane and ammonia at low pressure*, Vacuum **43**, 227–230 (1992).

[24] K. R. Williams and R. S. Muller, *Etch Rates for Micromachining Processing*, J. Microelectromech. S. **5**, 156–269 (1996).

[25] K. R. Williams, K. Gupta, and M. Wasilik, *Etch Rates for Micromachining Processing—Part II*, J. Microelectromech. S. **12**, 761–778 (2003).

[26] M. J. Dykstra and L. E. Reuss, *Biological Electron Microscopy; Theory, Techniques, and Troubleshooting*, Springer Netherlands (2003).

[27] P. W. Hawkes and E. Kasper, *Principles of Electron Optics*, vol. 2, Academic Press, London (1996).

[28] R. Gomer, *Field Emission and Field Ionization*, Harvard University Press (1961).

[29] A. Degiovanni and R. Morin, *Low Energy Electron Interferences using a Biprism-projection Microscope Combination*, in *Proceedings of the ICEM 13*, vol. 1, pp. 331–332, ICEM, Paris (1994).

[30] O. Klemperer, *Electron Optics*, Cambridge University Press (1953).

[31] M. Oral and B. Lencová, *Calculation of aberration coefficients by ray tracing*, Ultramicroscopy **109**, 1365–1373 (2009).

[32] P. W. Hawkes and E. Kasper, *Principles of Electron Optics*, vol. 1, Academic Press, London (1996).

[33] M. Born and E. Wolf, *Principles of Optics*, Cambridge University Press (1999).

[34] D. Sayre, *Some implications of a theorem due to Shannon*, Acta Cryst. **5**, 843 (1952).

[35] R. Gerchberg and W. Saxton, *A Practical Algorithm for the Determination of Phase from Image and Diffraction Plane Pictures*, Optik **35**, 237–246 (1972).

[36] J. Fienup, *Reconstruction of an object from modulus of its Fourier-transform*, Opt. Lett. **3**, 27–29 (1978).

[37] J. Miao, P. Charalambous, J. Kirz, and D. Sayre, *Extending the methodology of X-ray crystallography to allow imaging of micrometre-sized non-crystalline specimens*, Nature **400**, 342–344 (1999).

[38] J. Miao et al., *High Resolution 3D X-Ray Diffraction Microscopy*, Phys. Rev. Lett. **89**, 088303 (2002).

[39] I. K. Robinson, I. A. Vartanyants, G. J. Williams, M. A. Pfeifer, and J. A. Pitney, *Reconstruction of the Shapes of Gold Nanocrystals Using Coherent X-ray Diffraction*, Phys. Rev. Lett. **87**, 195505 (2001).

[40] G. J. Williams, M. A. Pfeifer, I. A. Vartanyants, and I. K. Robinson, *Three-Dimensional Imaging of Microstructructure in Au Nanocrystals*, Phys. Rev. Lett. **90**, 175501 (2003).

[41] H. He et al., *Experimental lensless soft-X-ray imaging using iterative algorithms: phasing diffuse scattering*, Acta Cryst. A **59**, 143–152 (2003).

[42] J. Miao et al., *Imaging whole Escherichia coli bacteria by using single-particle x-ray diffraction*, PNAS **100**, 110–112 (2003).

[43] D. Shapiro et al., *Biological imaging by soft X-ray diffraction microscopy*, PNAS **102**, 15343–15346 (2005).

[44] H. N. Chapman et al., *High-resolution* ab initio *three-dimensional x-ray diffraction microscopy*, J. Opt. Soc. Am. A **23**, 1179–1200 (2006).

[45] J. Zuo, I. Vartanyants, M. Gao, R. Zhang, and L. Nagahara, *Atomic Resolution Imaging of a Carbon Nanotube from Diffraction Intensities*, Science **300**, 1419–1421 (2003).

[46] S. Morishita, J. Yamasaki, K. Nakamura, T. Kato, and N. Tanaka, *Diffractive imaging of the dumbbell structure in silicon by spherical-aberration-corrected electron diffraction*, Appl. Phys. Lett. **93**, 183103 (2008).

[47] O. Kamimura, T. Dobashi, K. Kawahara, T. Abe, and K. Gohara, *10-kV diffractive imaging using newly developed electron diffraction microscope*, Ultramicroscopy **110**, 130–133 (2010).

[48] J. Miao, K. Hodgson, and D. Sayre, *An approach to three-dimensional structures of biomolecules by using single-molecule diffraction images*, PNAS **98**, 6641–6645 (2001).

[49] H. N. Chapman et al., *Femtosecond diffractive imaging with a soft-X-ray free-electron laser*, Nat. Phys. **2**, 839–843 (2006).

[50] I. Müllerová and L. Frank, *Scanning Low-Energy Electron Microscopy*, Adv. Imag. Elect. Phys. **128**, 309–443 (2003).

[51] L. Reimer, *Scanning Electron Microscopy*, Springer Verlag, Berlin, Heidelberg (1985).

[52] H. J. Fitting, E. Schreiber, Kuhr, and A. von Czarnowski, *Attenuation and escape depths of low-energy electron emission*, J. Electron Spectrosc. Relat. Phenom. **119**, 35–47 (2001).

[53] B. Boudaïffa, P. Cloutier, D. Hunting, M. A. Huels, and L. Sanche, *Resonant Formation of DNA Strand Breaks by Low-Energy (3 to 20eV) Electrons*, Science **287**, 1658–1660 (2000).

[54] J. M. Cowley, *Diffraction Physics*, Elsevier Science B. V. , Amsterdam (1995).

[55] E. Hecht, *Optics*, Addison Wesley, San Francisco (2002).

[56] J. H. C. Spence, U. Weierstall, and M. Howells, *Phase recovery and lensless imaging by iterative methods in optical, X-ray and electron diffraction*, Phil. Trans. R. Soc. Lond. A **360**, 875–895 (2002).

[57] R. Millane, *Phase retrieval in crystallography and optics*, J. Opt. Soc. Am. A **7**, 394–411 (1990).

[58] Y. M. Bruck and L. G. Sodin, *On the Ambiguity of the Image Reconstruction Problem*, Opt. Commun. **30**, 304–308 (1979).

[59] R. Barakat and G. Newsam, *Necessary conditions for a unique solution to two-dimensional phase recovery*, J. Math. Phys. **25**, 3190–3193 (1984).

[60] R. H. T. Bates, *Fourier phase problems are uniquely solvable in more than one dimension. I: Underlying theory*, Optik **61**, 247–262 (1982).

[61] R. G. Lane, W. R. Fright, and R. H. T. Bates, *Direct Phase Retrieval*, IEEE T. Acoust. Speech **35**, 520–526 (1987).

[62] A. Walther, *The question of phase retrieval in optics*, J. Mod. Optic. **10**, 41–49 (1962).

[63] J. Miao, D. Sayre, and H. Chapman, *Phase retrieval from the magnitude of the Fourier transforms of nonperiodic objects*, J. Opt. Soc. Am. A **15**, 1662–1669 (1998).

[64] R. P. Millane, *Multidimensional phase problems*, J. Opt. Soc. Am. A **13**, 725–734 (1996).

[65] J. Fienup, *Phase retrieval algorithms – a comparison*, Appl. Optics **21**, 2758–2769 (1982).

[66] K. S. Raines et al., *Three-dimensional structure determination from a single view*, Nature **463**, 214–217 (2010).

[67] R. W. Gerchberg and W. O. Saxton, *Phase Determination from Image and Diffraction Plane Pictures in the Electron Microscope*, Optik **34**, 275–284 (1971).

[68] J. Fienup, *Reconstruction of a complex-valued object from the modulus of its Fourier-transform using a support constraint*, J. Opt. Soc. Am. A **4**, 118–123 (1987).

[69] S. Marchesini et al., *X-ray image reconstruction from a diffraction pattern alone*, Phys. Rev. B **68**, 140101 (2003).

[70] J. S. Wu, U. Weierstall, J. C. H. Spence, and C. T. Koch, *Iterative phase retrieval without support*, Opt. Lett. **29**, 2737–2739 (2004).

[71] W. Huang, B. Jiang, R. Sun, and J. Zuo, *Towards sub-Å atomic resolution electron diffraction imaging of metallic nanoclusters: A simulation study of experimental parameters and reconstruction algorithms*, Ultramicroscopy **107**, 1159–1170 (2007).

[72] J. S. Wu and J. C. H. Spence, *Reconstruction of complex single-particle images using charge-flipping algorithm*, Acta Cryst. A **61**, 194–200 (2005).

[73] G. Oszlányi and A. Sütő, Ab initio *structure solution by charge flipping*, Acta Cryst. A **60**, 134–141 (2004).

[74] J. Miao, T. Ishikawa, E. Anderson, and K. Hodgson, *Phase retrieval of diffraction patterns from noncrystalline samples using the oversampling method*, Phys. Rev. B **67**, 174104 (2003).

[75] F. Wicki, *Kohärente Beugung niederenergetischer Elektronen: Beamstop und Probenpräparation*, Bachelor Thesis (2010).

[76] P. Thibault, *Feasibility of 3D reconstructions from a single 2D diffraction measurement*, arXiv:0909.1643v2 [physics.data-an].

[77] J. Miao and D. Sayre, *On possible extensions of X-ray crystallography through diffraction-pattern oversampling*, Acta Cryst. A **56**, 596–605 (2000).

[78] J. Spence, U. Weierstall, and M. Howells, *Coherence and sampling requirements for diffractive imaging*, Ultramicroscopy **101**, 149–152 (2004).

[79] D. Bohm, *Quantum Theory*, Prentice Hall, New York (1951).

[80] R. H. Fowler and L. W. Nordheim, *Electron Emission in Intense Electric Fields*, Proc. Roy. Soc. Lond. A **119**, 173–181 (1928).

[81] T. E. Stern, B. S. Gossling, and R. H. Fowler, *Further Studies in the Emission of Electrons from Cold Metals*, Proc. Roy. Soc. Lond. A **124**, 699–723 (1929).

[82] L. W. Nordheim, *The Effect of the Image Force on the Emission and Reflexion of Electrons by Metals*, Proc. Roy. Soc. Lond. A **121**, 626–639 (1928).

[83] R. E. Burgess, H. Kroemer, and J. M. Houston, *Corrected Values of Fowler-Nordheim Field Emission Functions $\theta(y)$ and $s(y)$*, Phys. Rev. **90**, 515 (1953).

[84] R. D. Young, *Theoretical Total-Energy Distribution of Field-Emitted Electrons*, Phys. Rev. **113**, 110–114 (1959).

[85] R. D. Young and E. W. Müller, *Experimental Measurement of the Total-Energy Distribution of Field-Emitted Electrons*, Phys. Rev. **113**, 115–120 (1959).

[86] E. Ruska, *Zur Fokussierbarkeit von Kathodenstrahlbündeln großer Ausgangsquerschnitte*, Z. Phys. **83**, 684–697 (1933).

[87] H. Liebl, *The image aberration caused by the acceleration field between concentric spherical electrodes*, Optik **83**, 129–133 (1989).

[88] H.-W. Fink and H. Schmid, *Atomic Resolution in Lensless Low-Energy Electron Holography*, Phys. Rev. Lett. **67**, 1543–1546 (1991).

[89] H.-W. Fink, W. Stocker, and H. Schmid, *Holography with Low-Energy Electrons*, Phys. Rev. Lett. **65**, 1204–1206 (1990).

[90] J. C. H. Spence, Q. Qian, and M. P. Silverman, *Electron source brightness and degeneracy from Fresnel fringes in field emission point projection microscopy*, J. Vac. Sci. Technol. A **12**, 542–547 (1994).

[91] J. C. H. Spence and W. Qian, *Transmission-electron Fourier imaging of crystal lattices using low-voltage field-emission sources: Theory*, Phys. Rev. B **45**, 10271–10279 (1992).

[92] R. Morin, *Point source physics: Application to electron projection microscopy and holography"*, Microsc. Microanal. M. **5**, 501–508 (1994).

[93] A. P. Janssen and J. P. Jones, *The sharpening of field emitter tips by ion sputtering*, J. Phys. D: Appl. Phys. **4**, 118–123 (1971).

[94] COMSOL Multiphysics 3.5a, 1998-2008 COMSOL AB.

[95] H. Liebl, *Applied Charged Particle Optics*, Springer Berlin Heidelberg (2008).

[96] SIMION 8.0, Scientific Instrument Services, Inc., 2003–2006.

[97] T. H. Hoenderken, C. W. Hagen, J. E. Barth, P. Kruit, and G. O. Nützel, *Influence of the microchannel plate and anode gap on the spatial resolution of an image intensifier*, J. Vac. Sci. Technol. B **19**, 843–850 (2001).

[98] P. Thibault and I. C. Rankenburg, *Optical diffraction microscopy in a teaching laboratory*, Am. J. Phys. **75**, 827–832 (2007).

[99] W. H. Dobelle and M. Beer, *Chemically Cleaved Graphite Support Films for Electron Microscopy*, J. Cell Biol. **39**, 733–736 (1968).

[100] J. C. Meyer et al., *The structure of suspended graphene sheets*, Nature **446**, 60–63 (2007).

[101] J. C. Meyer, C. O. Girit, M. F. Crommie, and A. Zettl, *Imaging and dynamics of light atoms and molecules on graphene*, Nature **454**, 319–322 (2008).

[102] K. S. Novoselov et al., *Two-dimensional atomic crystals*, PNAS **102**, 10451–10453 (2005).

[103] Z. Wei, D. E. Barlow, and P. E. Sheehan, *The Assembly of Single-Layer Graphene Oxide and Graphene Using Molecular Templates*, Nano Lett. **8**, 3141–3145 (2008).

[104] J. I. Paredes, S. Villar-Rodil, A. Martínez-Alonso, and J. M. D. Tascón, *Graphene Oxide Dispersions in Organic Solvents*, Langmuir **24**, 10560–10564 (2008).

[105] R. S. Pantelic, J. C. Meyer, U. Kaiser, W. Baumeister, and J. M. Plitzko, *Graphene oxide: A substrate for optimizing preparations of frozen-hydrated samples*, J. Struc. Biol. **170**, 152–156 (2010).

[106] N. R. Wilson et al., *Graphene Oxide: Structural Analysis and Application as Highly Transparent Support for Electron Microscopy*, ACS Nano **3**, 2547–2556 (2010).

[107] I. Müllerová, M. Hovorka, R. Hanzlíková, and L. Frank, *Very Low Energy Scanning Electron Microscopy of Free-Standing Ultrathin Films*, Mater. Trans. **51**, 265–270 (2010).

[108] J. C. Meyer, M. Paillet, G. S. Duesberg, and S. Roth, *Electron diffraction analysis of individual single-walled carbon nanotubes*, Ultramicroscopy **106**, 176–190 (2006).

[109] R. J. Chen et al., *Noncovalent functionalization of carbon nanotubes for highly specific biosensors*, PNAS **100**, 4984–4989 (2003).

[110] K. Bradley, M. Briman, A. Star, and G. Grüner, *Charge Transfer from Adsorbed Proteins*, Nano Lett. **4**, 253–256 (2004).

[111] S. C. Tsang et al., *Immobilization of Small Proteins in Carbon Nanotubes: High-resolution Transmission Electron Microscopy Study and Catalytic Activity*, J. Chem. Soc., Chem. Commun. pp. 1803–1804 (1995).

[112] R. J. Chen, Y. Zhang, D. Wang, and H. Dai, *Noncovalent Sidewall Functionalization of Single-Walled Carbon Nanotubes for Protein Immobilization*, J. Am. Chem. Soc. **123**, 3838–3839 (2001).

[113] W. Huang et al., *Attaching Proteins to Carbon Nanotubes via Diimide-Activated Amidation*, Nano Lett. **2**, 311–314 (2002).

[114] S. Bhattacharyya, C. Sinturel, J. P. Salvetat, and M.-L. Saboungi, *Protein-functionalized carbon nanotube-polymer composites*, Appl. Phys. Lett. **86**, 113104 (2005).

[115] K. Jiang et al., *Protein immobilization on carbon nanotubes vie a two-step process of diimide-activated amidation*, J. Mater. Chem. **14**, 37–39 (2004).

[116] A. Hirsch, *Functionalization of Single-Walled Carbon Nanotubes*, Angew. Chem., Int. Ed. **41**, 1853–1859 (2002).

[117] J. Meyer, *Structure and Properties of Carbon Nanotubes*, Ph.D. thesis, Universität Tübingen (2006).

[118] Y. Li et al., *Growth of Single-Walled Carbon Nanotubes from Discrete Catalytic Nanoparticles of Various Sizes*, The Journal of Physical Chemistry B **105**, 11424–11431 (2001).

[119] J. Frank, *Sinlge-Particle Imaging of Macromolecules by Cryo-Electron Microscopy*, Annu. Rev. Biophys. Biomol. Struct. **31**, 303–319 (2001).

[120] S. P. Hau-Riege, R. A. London, H. N. Chapman, A. Szoke, and N. Timneanu, *Encapsulation and Diffraction-Pattern-Correction Methods to Reduce the Effect of Damage in X-Ray Diffraction Imaging of Single Biological Molecules*, Phys. Rev. Lett. **98**, 198302 (2007).

[121] J. C. H. Spence and P. W. Hawkes, *Diffract-and-destroy: Can X-ray lasers "solve" the radiation damage problem?*, Ultramicroscopy **108**, 1502–1503 (2008).

Mein herzlicher Dank gilt Hans-Werner Fink, Conrad Escher,
Tatiana Latychevskaia und Jean-Nicolas Longchamp
für ihre Unterstützung im Zusammenhang mit der vorliegenden Arbeit.

Curriculum Vitae

Elvira Steinwand

Date of birth: 09.12.1981
Place of birth: Tübingen
Citizenship: German

Education

July 2000	University-entrance diploma (Abitur), Droste-Hülshoff-Gymnasium Meersburg, Germany
04/2001–08/2003	Studies of physics, University of Heidelberg, Germany subsidiary subjects: mathematics, chemistry
08/2003–02/2004	Erasmus-scholarship at the University of Santiago de Compostela, Spain
02/2004–09/2005	Studies of physics, University of Heidelberg, Germany elective courses: computational physics, biophysics subsidiary subjects: mathematics
09/2005–09/2006	Diploma thesis in physics *4Pi Mikroskopie - Anwendung und Analyse bildverarbeitender Methoden* at the Kirchhoff Institute for Physics, University of Heidelberg
09/2006	Diploma degree in physics, University of Heidelberg
since Dec 2006	PhD studies on *Low-energy electron coherent diffraction imaging* in the group of Prof. Dr. H.-W. Fink at the Physics Institute of the University of Zurich, Switzerland

---

Teaching

2002-2003	Teaching assistant, Mathematical department, University of Heidelberg, undergraduate courses for students of mathematics and physics
since Dec 2006	Teaching assistant, Physics Institute, University of Zurich, basic courses and lab-courses for medical and for physics students

PUBLICATIONS

E. Steinwand, J.-N. Longchamp, H.-W. Fink, *Fabrication and characterization of low-aberration micrometer sized electron lenses*, Ultramicroscopy 110 (2010) 1148–1153

E. Steinwand, J.-N. Longchamp, H.-W. Fink, *Coherent low-energy electron diffraction of individual nanometer sized objects*, submitted

TALKS AND POSTER PRESENTATIONS

Coherent Diffraction with Low Energy Electrons, Workshop on *Current Trends in Structural Biology on the Single Molecule Level*, Lucerne, January 26 - 29, 2009

Coherent low-energy electron diffraction microscopy of single bio-molecules, poster presentation, International Conference on *Nanostructures at Surfaces*, Ascona, September 20 - 25, 2009

Coherent Diffraction Microscopy with Low Energy Electrons, poster presentation, Workshop, Swiss Soft Days, Zurich, February 4, 2010

Low-energy electron coherent diffraction microscopy, poster presentation, Doctoral School in Biophysics, Crans Montana, February 14 - 20, 2010

Coherent diffraction microscopy with low-energy electrons, International Conference *From Solid State to Biophysics*, Dubrovnik, June 12 - 19, 2010

Coherent diffraction imaging with low-energy electrons, Workshop, Swiss Soft Days, Fribourg, October 20, 2010